The Role of Trading Intensity in Duration Modelling and Price Discovery

Evidence from the European Carbon market

Volume I

Doctoral Thesis Submitted by

Kalaitzoglou Iordanis Angelos

For the completion of a three years PhD program

July 2011

Heriot-Watt University
School of Management and Languages
Department of Accountancy, Economics and Finance
Supervisors: **Mohamed Sherif**
Boulis Ibrahim
Julian Fennema

Abstract

In this study, trading intensity is employed to investigate the role of information and liquidity in duration modelling and price discovery in the two largest exchanges of the European Carbon market, namely European Climate Exchange (ECX) and Nord Pool (NP). First, duration modelling is examined for the first time in this market, and existing ACD models are empirically extended to explore the impact of stylized facts, such as non-linear effects of trading intensity and OTC transactions. Second, the “time dimension” of information is investigated focusing on the informational content of trading intensity. A Smooth-Transition-Mixture of Weibull Distributions ACD (STM-ACD) model that distinguishes between three types of trades is proposed. Time, volume and OTC transactions measure how related related a trade is to information. Third, the price impact of the “time dimension” of information is examined. A new dynamic expectations, structural pricing model is proposed in order to account for the learning process of traders and their expectations. Trading intensity is used to measure the sensitivity of market participants to information and liquidity.

The main findings indicate that empirical adjustments significantly improve duration modelling. In consistence with Bauwens et al. (2004), the specification of the conditional mean contributes more to model performance. Trading intensity appears to create a momentum, especially in ECX, whereas OTC transactions seem to slow down the trading process, probably due to information inflow, especially in NP. Furthermore, similar to Easley and O’Hara (1992) higher trading intensity is associated with increased presence of information. Trading intensity is found to be able to distinguish among three different types of trades, according to their informational content. The timing of acquiring information can make it further exploitable. A significant proportion of uninformed traders in the Carbon market is found to observe the market trying to extract price unresolved information. Consequently, informed traders are found to act strategically, according to Kyle (1985), but they are less efficient in covering their actions as market gains complexity, mainly because of higher liquidity levels and improved learning process. In addition, large transactions appear to increase the information price component, while the liquidity component seems to asymmetrically decrease, probably due to economies of scale. Consequently, trading intensity appears to have a dual impact on price, spread and price change volatility, which is determined by current market conditions and dealers’ exposure to risk. Finally, market making in this market seems to be profitable only when expected trading intensity is low.

To Marianthi

for the countless hours of patience

and

to my unborn child

because

“nothing else matters”

Acknowledgements

The present study would have never been completed if it were not for the financial support and the research environment, generously provided by Heriot-Watt University. In addition, special thanks should be given to Mr. Pontus Ripstrand from Nord Pool and Mrs Stefanie Muller from European Climate Exchange for their co-operation in collecting the data needed.

I would also like to personally thank my first two PhD thesis supervisors. Dr. Mohamed Sherif for supporting my application for the school's scholarship, in the early stages of my studies, and for his continuous trust, and guidance in my teaching activities. He gave me countless opportunities for enforcing my academic position, for which I am grateful. In addition, I would like to specially thank Dr. Boulis Ibrahim, who has been an academic inspiration, with his comments and guidance, which were immeasurably valuable. Especially our endless meetings acted as a catalyst that changed dramatically my academic way of thinking. This is an asset I will carry in my academic life.

Further, I would like to thank Mrs Marianthi Themeli, Dr Adele Thomas and Dr Panos Vlachopoulos for their invaluable help and comments in the editing my manuscript. Finally, I would like to specially thank Mrs Lori Heaford for proofreading my thesis.

ACADEMIC REGISTRY

Research Thesis Submission

Name:			
School/PGI:			
Version: *(i.e., First, Resubmission, Final)*		Degree Sought (Award **and** Subject area)	

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1) the thesis embodies the results of my own work and has been composed by myself
2) where appropriate, I have made acknowledgement of the work of others and have made reference to work carried out in collaboration with other persons
3) the thesis is the correct version of the thesis for submission and is the same version as any electronic versions submitted*.
4) my thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
5) I understand that as a student of the University I am required to aBide by the Regulations of the University and to conform to its discipline.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate:		Date:	

Submission

Submitted By *(name in capitals):*	
Signature of Individual Submitting:	
Date Submitted:	

For Completion in the Student Service Centre (SSC)

Received in the SSC by *(name in capitals):*			
Method of Submission *(Handed in to SSC; posted through internal/external mail):*			
E-thesis Submitted (mandatory for final theses)			
Signature:		Date:	

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1. Introduction

Finance theory has developed considerably, over the last decades providing new insights into market structure and price formation. The availability of Ultra-high-Frequency (UHF) data offers the opportunity to examine the trading process on a whole new level. All transactions are recorded, and individual events, such as profit announcements and intraday phenomena, can be examined separately. This provides researchers with the opportunity to use all available market information, with the same frequency as market participants. This allows the study of issues such as how traders aggregate information, how they interpret it and how they act upon receiving it. In addition, trading characteristics, such as trading frequency and volume, might have an impact on subsequent trading and price formation.

This can be particularly important in new markets with unique characteristics, such as the Carbon market. The analysis of intraday market dynamics and their price impact can improve trading practice and, therefore, market quality. The understanding of various behavioural aspects of intraday trading patterns can enhance flexibility in regulation, which can be adjusted according to particular market needs, as they arise in real time. This provides the foundation for more efficient pricing, by sustaining a “healthy” trading environment without sacrificing liquidity or information aggregation. The new market can develop and gain complexity and maturity faster, providing more sophisticated environment and financial instruments that can reflect more accurately market “fundamentals”. This can help the Carbon market to service its purpose, to reduce emissions, more effectively.

1.1 Intraday Price Formation

O’Hara (1995) maintains that, on intraday level, prices are generally seen as being determined by the actions of market participants, and not as a macroeconomic phenomenon. Their access to relevant information and their liquidity desires, which are seen as the main price determinants, determine their strategies. They observe the market, they gather information and they build their trading strategies combining new information with liquidity needs, which are derived from their optimal portfolio position and potential deviations from it. Price equilibrium is achieved as a balance between various individual trading approaches. Clearly, time and size dimensions rise, since not all market participants’ needs coincide in time or size.

Early literature emphasizes the price impact of information and liquidity. Unlike Stigler (1967) who sees trading costs as market imperfections, Demsetz (1968) and Bagehot (1971) argue that this is a natural outcome of the trading process. Dealers face a continuous matching problem and they try to compensate for any severe order flow imbalances, by quoting a different price for selling (Ask) and buying (Bid). Consequently, the Bid-Ask spread is seen as the price of "immediacy", which heavily depends on incoming information and liquidity levels, and the dealer's exposure to the market.[1] This changes current market prices and reveals spreads as being directly proportional to order flow, and can definitely no longer be ignored.

"Order processing"

Naturally, the literature expands to include investigations into the nature of these costs specifically to their composition as derived by trading activity. The first component identified, is a fixed cost applied to all transactions. This component is known as the order-processing cost and is formally recognized first by Demsetz (1971). As further studies, such as Stoll (1978), Huang and Stoll (1997) and Ahn et al. (2002), mention, this cost should be variable per unit as a decreasing function of volume, since it is fixed per transaction. Further, Bollen et al. (2004) recognize that this cost should be a decreasing function of competition as well. Dealers can no longer charge a large fixed component, because they might lose sales. Instead, order-processing cost should asymptotically converge, in perfect competition, to the marginal cost of providing liquidity.

"Inventory" Models

Different studies emphasize the liquidity effect as a main price determinant. In a new class of models, originating back to Tinic (1972), Tinic and West (1972) and Benston and Hagerman (1974), known as "inventory-holding" models, price setting is seen as being influenced directly by dealers' inventories. Since dealers need to continuously provide liquidity, they need to set their prices to face two types of risk. One is related to increasing inventory (i.e., excessive cost of carry, especially overnight where information might change significantly) while the other is associated with decreasing inventory (i.e., loss of sales). In the first case, when they keep high inventories,

[1] Market makers are introduced into markets to enhance liquidity. They commit themselves to continuously support the trading process, by being ready to buy or sell (i.e., providing "immediacy"), in exchange for various market privileges, such as real time full market information or reduced clearing or trading fees. For further information please refer to O'Hara (1995) and Jong and Rindi (2009).

information might change drastically over a short period of time, or overnight, and they might end up in an adverse position due to excessive carrying cost or price changes. On the second case, they might miss opportunities of increasing demand, because they do not possess the asset, and therefore they might need to buy it at spot, expensive, prices. Dealers continuously observe the market and adjust their quoted prices and inventories act as a buffer to trading activity that reflects current liquidity. Therefore, inventories do not change the perception of investors about the "fair" value of the underlying asset. Hasbrouck (1988, 1991a, 1991b) therefore concludes that their effect on prices is only temporary.

These models focus on the "viability" of market makers (Garman, 1976), on the maximization of their profits (Amihud and Mendelson, 1980), or the maximization of their terminal wealth (Stoll, 1978; Ho and Stoll, 1981, 1983). Several other studies, such as Cohen et al. (1980, 1981), O'Hara and Oldfield (1986), Angel (1994) and Harris (1998), generalize investigating potential inventory-holding effects on hybrid markets, where market as well as limit orders can be submitted. This introduces another source of uncertainty, uncertainty of execution. Cohen et al. (1980, 1981) develop the idea of "gravitational pool", where the aggressiveness of market participants determines the appropriate type, time and size of the order.[2]

Inventory models, although they present noticeably different approaches, raise four points that are relevant to the present study. First and most important, these models recognize the temporary price impact of liquidity. Second, they recognize that market participants observe order flow to formulate expectations concerning future levels of liquidity, focusing on imbalances of incoming order flow. Consequently, they recognize the importance of intensity of trading as influential to price formation. Volume of trading plays an important role in the matching of orders and it might move market makers from their optimal positions. Market makers, therefore, alter their quotations to deal with incoming imbalances. Third, these imbalances occur because of the asynchronous arrival of orders of different initiation, Buyer or Seller. Time is implicitly

[2] The issue of Limit versus Market orders and especially the information efficiency of these markets (Brown and Whang, 1997), the relation between type of orders and type of traders (O'Hara and Oldfield, 1986; Chadravarty and Holden, 1995) and the optimal choice between them, given current market conditions (see, inter alia, Cohen et al., 1981; Angel, 1994; Kumar and Seppi, 1994; Harris and Hasbrouck, 1996; Parlour, 1998; Foucault, 1999; Goettler et al., 2005; Wald and Horrigan, 2005), has been discussed extensively in the literature. More recently Bouchaud et al. (2004, 2006) maintain that the marginal profit of the type of order, either in the form of capital gains or in the form of Bid-Ask spread, minimizes arbitrage opportunities by allowing market participants to submit either of them. Whichever the case, the vast majority of studies confirms that limit orders are associated with wider spreads.

assumed to have an impact on intraday price formation, although order flow is still the main source of uncertainty. Finally, the price impact of different market structures is further examined. In particular, the significance of the choice of which type of order to submit is emphasized, introducing a new element of risk.

"Information" Models

In parallel, another class of models has been developed, underlying the role of information as an intraday price determinant. In a seminal study, Bagehot (1971) argues that spreads would exist even in the absence of explicit trading costs. He explains that when market makers transact with better informed traders, their trading will on average incur losses. However, they can compensate their "immediacy" and "lack of knowledge" by transacting with uninformed traders. They do so by charging a fee, in the form of Bid-ask spread, which should be wider when they believe that the presence of information based trading is increased. This idea has been developed further in "information models", and emphasizes the informational content of past trades and how this information is incorporated into prices.

These models see intraday price formation, as an equilibrium price arising by a "game" among different market participants. Exogenous information is assumed to arrive randomly and these market participants are grouped according to their access to unresolved, price-relevant information in terms of both level and time dimensions. Early literature (cf., Kyle, 1985; Glosten and Milgrom, 1985) recognizes two types of traders. Market participants in the first group, namely uninformed, are assumed to trade for reasons unrelated to the arrival of new information, and they do so because of their individual liquidity needs. In contrast, informed traders possess information, which will be incorporated into subsequent prices, before this information is revealed to other market participants. Therefore they have a timing advantage over the rest of the market and have an incentive to exploit it, at the expense of market makers and uninformed traders. However, subsequent studies, such as Admati and Pfleiderer (1988), recognize that the assumption of the "naivety" of uninformed traders is too strong and is not applicable in real markets. Instead, they recognize that uninformed traders can be further divided according to whether they can chose the time and size of their transaction in order to maximize their "terminal wealth".[3] Admati and Pfleiderer (1988)

[3] This term is borrowed by economics theory and it describes an essential motivation for this group of traders to engage themselves into a transaction. If they do not have that incentive, there is no reason to

distinguish between non-discretionary, naive uninformed traders, who trade only for liquidity reasons, and discretionary liquidity traders, who observe the market to extract information upon which they can act.

Two main streams have been developed, according to how informed traders exploit their informational "privilege". First, in "sequential" models (Glosten and Milgrom, 1985; Easley and O'Hara, 1987, 1992; Easley et al., 1997, 2002), trades occur in a sequence and informed traders act at once, trading in high volumes, as much as the current market conditions and their trading capacity allows. They do not take into account the price effect of their trades, because they do not return to the market before the information they possess is fully resolved into prices. Market makers observe past transactions to extract price-relevant information and to formulate a Probability of INformed (PIN) trading, upon which they base their price quotations. In contrast, according to "strategic" models (Kyle, 1985; Kyle, 1989; Foster and Viswanathan, 1993; Holden and Subrahmayam, 1992; Back 1992) informed traders return to the market before information is fully resolved. Convergence to fully informed equilibrium cannot occur immediately after the transaction and, therefore, informed traders need to take into account the post-trade effect of their actions. They need to act strategically, by segmenting their trades, in order to fully exploit their information advantage.

Information models raise the importance of four issues. First, both recognize different types of trades/traders and describe intraday price formation as an outcome, equilibrium, of the interaction of these traders. Second, and particularly apparent in more recent studies, uninformed traders have an incentive to learn. These traders act according to expectations formulated from observations extracted from trading history. Third, they emphasize the importance of order flow and trading volume. In particular, trading time and size are important to identify PIN in sequential models. Finally, following Easley and O'Hara (1987, 1992) the literature recognizes the dual effect of trading intensity, and thus of past trading activity, on prices, although it is difficult to distinguish the permanent from the transitory component (c.f., Madhavan, 2000). Past transactions carry both an informational component that has a permanent price impact, because it revises beliefs about the "fair" value of the asset, and a liquidity component, which has only a temporary influence on prices because of expected order flow variations.

---

enter the market and no equilibrium can be achieved. Similarly, with inventory models, in that case, the market fails. See De Jong and Rindi (2009).

"Trade Indicator" Models

The resilience of both information and inventory models to "game" theory gives them a theoretical character, and, therefore, their empirical application is fairly limited.[4] In contrast, another branch of literature focuses on empirical issues and examines intraday price formation at transaction level. Roll (1984) proposes an empirical way to extract trading spreads, summarized in one component, from covariances of price changes. Though simplistic, his approach provides a convenient way to model prices and estimate spreads. Choi et al. (1988), Stoll (1989) and George et al. (1991) extend Roll's (1984) models, incorporating all three price components. In addition, Hasbrouck (2007) suggests that, with certain transformations and assumptions for the innovations, Roll's (1984) model could be transformed into a VAR process.[5]

Glosten and Harris (1998), introduce another source of uncertainty for both the informational and liquidity component of trades. A trade initiation variable, is allowed to drive both price determinants. Similar to inventory models, price quoting depends on the trade initiation, and since market makers cannot forecast the direction of the next trade, they quote two prices (Bid and Ask) and therefore spreads can be estimated. Ho and Stoll (1997) propose a general model that nests Autoregressive models and other structural approaches as special cases. Their model is "complete" in the sense that all three cost components can be derived. However, estimates of half spreads are required prior to cost component computation. Madhavan et al. (1997) propose a similar approach, in which the information and order-processing components can be estimated in one step, where the innovation in order flow measures price-relevant information.

1.2 Further Issues

1.2.1 Trading Volume

Microstructure literature raises several issues, recognizing their potential, direct or indirect, impact on price formation. One of the variables that attracted the early interest of researchers is trading size. Even in the inventory model developed by Stoll (1978), the relation between trading volume and spreads follows a U-shape pattern. In contrast,

[4] O'Hara (1995) comments that for these models to be applicable, the "game" (i.e., the market structure, market participants, etc.) must be known in advance. Consequently, these models cannot be the primary choice when empirical issues are of interest.
[5] Roll's (1984) model is a structural approach that describes a structure of price formation, while VAR modelling provides an analytical tool for investigating the price impact of a trade over time, with a view mainly at forecasting. Both should provide the same results under specific circumstances.

De Jong et al. (1995), Huang and Stoll (1997) and Ahn et al. (2002) report a decreasing effect of trading volume on spreads. More recently Angelidis and Benos (2009) confirm the decreasing relation between trading volume and spread, but only for the liquidity component of price formation. They attribute this effect to economies of scale and report an increasing information component for larger transactions. This is fully consistent with Easley and O'Hara's (1987, 1992) proposition that increased volume indicates higher presence of information-based trading. Similarly, Chan and Lakonishok (1995) further develop this idea that trading frequency increases after large transactions.

1.2.2 Transaction Time

Another aspect that has been extensively discussed in the literature is the price impact of time. Both inventory and information models implicitly incorporate a time dimension in price equilibria. In inventory models, dealers face a matching problem between incoming order flow imbalances and their inventories due to the asynchronous arrival of order submissions.[6] Information models focus on the realization of informed trades and on any post-trade effects, focusing on both trading size and time.[7] The first studies that explicitly incorporate time into microstructure analysis, examine the relation between duration and information. Diamond and Verrecchia (1987) show that, under short selling restrictions, informed traders are not able to trade upon the arrival of bad news. They therefore, associate long durations with bad news. In contrast, Easley and O'Hara (1992) conjecture that informed traders will always transact upon the arrival of new information. Consequently a low trading frequency should be a sign of "no news". Dufour and Engle (2000b) argue that increased trading frequency indicates increased presence of informed traders and thus increased price impact. Ben Sita (2010) distinguishes the permanent from the transitory effect of time, which are both found to increase the price impact of a trade. In contrast, Grammig et al. (2007) provide evidence of higher price revisions after stages of inactivity.

ACD

However, unlike data sets of lower frequency, in UHF data time is irregularly spaced. Consequently, basic model assumptions might be violated and well-known econometric models might not hold unconditionally. In a seminal study, Engle and Russell (1998)

[6] Even if the probability of a Buy or Sell is 50 percent, orders might arrive at different points in time and significant temporary imbalances, excessive number of Buys or Sells, might be created.
[7] Informed traders poses private information and they have the incentive to exploit it before it becomes public. In addition, this piece of information is revealed by their actions and a price revision follows.

introduce the Autoregressive Conditional Duration (ACD) framework to model time. [8] They argue that duration (i.e., time between two consecutive trades) measures economic time. Unlike calendar time, it measures trading activity, which might flow faster or slower, and it might carry price relevant information. They model duration as a stochastic dependent process and formulate actual durations as being derived from past realization times. Their model exhibits strong similarities to the GARCH literature, and ACD models require specification of the conditional mean, as a function of past durations and a density function with a positive support.[9]

The original model was criticized for the simplicity both. Bauwens and Giot (2000), Dufour and Engle (2000b), Zhang et al. (2001), Meitz and Teräsvirta (2006) and Jasiak (1998) challenge the linear assumption and propose non-linear specifications that model higher moments of duration and account for longer, persistent, memories. Another branch of the literature criticizes the deterministic character of ACD models and proposes a stochastic approach (c.f., Bauwens and Veredas, 2004; Ning, 2004; Strickland, 2006; Bauwens and Galli, 2009), generally referred to as Stochastic Conditional Duration (SCD).[10] In addition, several studies (e.g., Zhang et al, 2001; Grammig and Maurer, 2000; De Luca and Zuccolotto, 2006) argue that distributions with monotonic hazard functions, such as the Exponential and Weibull, are insufficient in capturing the data's higher moments and, therefore, fail to describe the Data Generation Process (DGP) of duration.

Furthermore, Bauwens et al. (2004) argue that a single distribution cannot capture the idiosyncrasy of duration series and that the main improvement comes from the conditional mean specification. Both statements are particularly relevant to the present study. First, a new class of models suggests that a mixture of distributions, modelled either with a Markov Switching framework (Hujer et al., 2002; Hujer and Vuletic,

---

[8] Engle and Russell (1998) provide evidence that duration (time) series are exogenous, highly persistent (i.e., they have long memory), over-dispersed (i.e., standard deviation is larger than the mean) and clustered (i.e., long (short) durations followed by long (short)).

[9] Various other factors could be introduced as well. However, the majority of studies emphasize the exogeneity assumption and investigate the Data Generation Process (DGP) of duration as being independent.

[10] In SCD models an innovation component is introduced into the conditional mean specification, that accounts for exogenous shocks. This approach, although it is more flexible, it recognizes that the DGP of duration is governed by two sources of uncertainty, that they need to be modelled jointly. This raises significant estimation difficulties and therefore these models have not gained popularity. However, Ghysels et al. (2004) argues that even SCD is restrictive because higher order conditional moments are ignored. They propose the Stochastic Volatility Duration (SVD), which is a joint modelling of the conditional mean and variance. They postulate that this approach is essential when intraday liquidity is of interest.

2007) or with switching regimes (De Luca and Zuccolotto, 2006) or using copulas (Wing, 2008; De Luca et al., 2008), should better describe the DGP of duration.[11] Second, Bauwens et al. (2004) suggest that empirical adjustments should improve the performance of the models. This could be beneficial in relatively new markets with distinct stylized facts, such as the Carbon market.

1.2.3 Asymmetric Information and Learning

In addition, previous literature connects duration with information, and mixture distribution ACD models have been employed to investigate this issue further. First, several studies (e.g., Wong et al., 2009; Tay et al.; 2009) investigate the relation between trading frequency, volume and the presence of informed trading, confirming Easley and O'Hara's (1992) propositions that higher trading activity indicates higher presence of informed traders. Second, Hujer and Vuletic (2007), drawing on information models, recognize that there are different types of traders in the market by their differing access to price-relevant information, namely informed and uninformed. Their trading activity should follow different trading patterns that can be described by different hazard functions of duration. Gerhard and Hautsch (2007) extend this idea, by allowing uninformed market participants to observe the market and extract information before it goes public.[12] This raises two important point that have been extensively discussed in the literature.

First, several studies focus on the nature of the observable information. Easley and O'Hara (1992) uses volume of trading, which is often employed to describe the intensity of trading. Later studies, such as Dufour and Engle (2000b), Bowe et al. (2007), Angelidis and Benos (2009) and Ben Sita (2010), distinguish between trading size and trading frequency. These two variables proxy trading activity and traders learn

[11] The Markov Switching and the switching regime frameworks differ significantly in how they treat modelling flexibility, although both pursue higher precision. In the Markov framework, the threshold variable is latent, while in the mixture of distributions it is pre-determined. This allows for greater flexibility in the determining of regimes, while restricts the characteristics of each regime. In contrast, the smooth transition framework employs an observable threshold variable, sacrificing generality, but allows data to determine the characteristics of each regime.

[12] The conditional intensity, or else the hazard function, of duration measures the probability of a transaction to occur, given that it has not occurred till now. This could proxy the trading rate of different market participants (groups) and could be used to distinguish among them. Consequently, the trading pattern of uninformed traders could be assumed to follow an Exponential distribution, since it has a flat hazard function. In contrast, informed traders should appear only upon the arrival of new information and therefore their trading rate should exhibit significant, sharp fluctuations over time. According to Gerhard and Hautsch (2000), traders who act fast upon the arrival of new information could be described by a decreasing hazard function, while for traders, such as portfolio managers, who need to aggregate information the probability of a transaction to occur should increase over time.

by observing past trades. Consequently, time and size are recognized to carry information signals, without necessarily including prior price changes. Easley and O'Hara (1992) and Dufour and Engle (2000b) relate the magnitude of these variables to the strength of information signals. High trading size and/or trading frequency have a higher price impact. Therefore traders can formulate expectations, concerning subsequent price changes, by observing current and historical trading activity.

Second, the learning process is further investigated. Although all traders observe the same stimulus, they do not necessarily react in the same way. This could depend on various reasons, ranging from varying liquidity desires to different optimal portfolio positions. One of the most extensively discussed reasons, however, is the traders' ability to extract information, or how and what they can learn from it. Several studies (Shiller, 1981, 1984; De Jong et al., 1990; Chamley, 2003) recognize the dynamic nature of the learning process in that traders observing the same signals might interpret these signals in various ways. This depends on their level of "prior knowledge", their "market share" or simply their "expectations".[13] Other studies (Townsend, 1978; Frydman, 1982; Blume and Easley, 1984, 1998; Feldman, 1978; Vives, 2008) emphasize that the only exploitable information is when traders learn about an equilibrium and not on an equilibrium.[14] These studies raise the importance of the time dimension in the learning process, where according to Vives (1993), Jun and Vives (1997) and Vives (2008), the speed of learning is a crucial determinant of trading strategies and price formation.

1.3 The Carbon market

Information asymmetry and speed of learning appear to be particularly relevant to the trading practice in relatively new markets, which are rather illiquid and have distinct stylized facts. The Carbon market is a fast growing market, which has gained significant complexity over the last years, and given the character of the market, information resolution appears to play an important role both in how informative prices are and the extent of liquidity levels. Viswanathan (2010) purports that the Carbon market needs

[13] Whatever the source of variation in beliefs is, traders might "herd" in the right or the wrong direction towards subsequent price movements (see, inter alia, Banerjee, 1992; Bikhchandani et al., 1992; Smith and Sorensen, 2000).

[14] According to empirical microstructure literature, uninformed traders learn any price-relevant information on an equilibrium, when prices include all relevant information. No further expectations can be exploitable, because there is no unresolved information. In contrast, discretionary liquidity traders can learn about an equilibrium before prices incorporate private information, by observing the market trying to identify information signals. These traders have the incentive to extract information, learn from it, formulate expectations and realize their strategies before the market reaches an equilibrium stage.

strict regulation in order to restrict manipulation. However, it should also allow for innovations of market participants that would enhance liquidity. A non-regulated, non-transparent market would be liquid, but inaccurate in terms of price. In contrast, a strictly regulated environment would increase price accuracy, but may decrease liquidity. Both results would defy the initial purpose of emissions' reduction.[15]

Furthermore, the importance of the Carbon market, in tackling Greenhouse Gases emissions is undeniable. However, the literature is sparse and the microstructure of the market has only recently been examined. Early studies focus on the nature (Kruger et al., 2007), the legal framework (Convery and Redmond, 2007), the design (Burtraw et al., 2002; Böhringer and Lange, 2005; Kosobud et al., 2005), the history and origins of the market (Mansanet-Bataller and Pardo, 2008), as well as the nature of various allowances (Chevallier, 2010).[16] Other studies (see, inter alia, Rubin, 1996; Schennach, 2000; Godal and Klaasen, 2006; Schleich et al., 2006; Daskalakis and Markelos, 2008; Daskalakis et al., 2009; Mansanet-Bataller and Pardo, 2007; Rotfuss et al., 2009) explore the price impact of banking restrictions or of National Allocation Plans (NAPs).[17] Other studies (Kosobud et al., 2005; Daskalakis et al., 2009) discuss the nature of the asset, focusing on whether it is a commodity or a financial product, mainly by employing a cost of carry model approach (inter alia, Uhrig-Homburg and Wagner, 2007; Truck et al., 2007; Daskalakis et al., 2006, 2009). Another branch of literature (Christiansen and Arvanitakis, 2005; Bunn and Fezzi, 2007; Alberola, 2007, 2008, 2009a, 2009b) investigates their relation to other commodities, reporting similar trends.

Another stream of literature (Benz and Truck, 2006; Paolella and Taschini, 2008; Borak et al., 2006) investigates the relation between price formation and volatility. Other studies (Chevallier, 2009; Vinocur, 2009; Rittler , 2009; Conrand et al., 2010) examine the same issue from an intraday perspective, where a bidirectional causality between

---

[15] Viswanathan (2010) is based on a seminal study of Kyle and Viswanathan (2008), who argue that organized markets improve the aggregation of diverse information, which, according to Hayek (1945), increases "market efficiency". However, this does not necessarily mean that it increases "price accuracy" as well. They explain that a highly manipulated market might be highly efficient, in terms of incorporating manipulation into price, reflecting current market forces. However, assets are not necessarily priced accurately, according to their fundamental macroeconomic characteristics. Consequently, market participants and their trading strategies influence asset prices, instead of "just reflecting" their "fair" values. In terms of Carbon pricing, over- or under-pricing would lead to a highly speculative environment, which would undermine the efforts of emissions' reduction.

[16] Emission allowances are the main financial instruments traded in the market. For more information please refer to Section 1.4.

[17] According to the regulatory framework applied on phase one and two and the transition between them, banking (storing allowances for future use) was not allowed. This turned the trading interest on more complex products like futures contracts. For further information on banking restrictions, phases of the market and NAPs, please refer to Section 1.4.

spot and futures is observed and the importance of NAPs is underlined. Vinocur (2009) postulates that traders under-react on new information, while Mansanet-Bataller et al. (2010) provide evidence that traders are influenced more by order flow than by Carbon-related information. This is relevant to Dufour and Engle (2000b) and Hasbrouck (1988, 1991a, 1991b), who argue that trading activity has a significant price impact. This impact seems to be stronger than exogenous information in the Carbon market. More recently, other studies explore the microstructure of the market, such as trading spreads (Frino et al., 2010), price leadership (Benz and Klar, 2008) and spreads between various types of compliance units (Mansanet-Bataller et al., 2010).

1.4 Market Description

This section provides a non-exhaustive introductory description of the foundations and stylized characteristics of the European Carbon market to support the analysis in the following chapters.[18] Section 1.4.1 presents briefly the “Kyoto Protocol” and the three “flexibility mechanisms”, which aim at emissions’ reduction. Section 1.4.2 focuses on the European action towards climate change, which is the organized market for Carbon Allowances, namely the European Union Emissions Trading System (EU ETS).

1.4.1 Kyoto Protocol

Understanding the undeniably negative impact of Greenhouse Gases (GG), the vast majority of countries has ratified a treaty known as the “Kyoto Protocol”, aiming at their emissions reduction.[19] More precisely, the Kyoto Protocol aims at the “*stabilization of GG concentration in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system*” (Article 2 UN FCCC, 1997). The agreement defines three periods, in which, countries, should meet their targets. Phase I (from 2005 to 2007) is the pilot period, Phase II (from 2008 to 2012) is the commitment period and Phase III (from 2013 to 2020) is the post-commitment period for re-evaluation and further adjustments. The Kyoto Protocol establishes three “flexibility mechanisms” that aim at diminishing energy-related costs, while achieving

[18]For an in depth analysis please refer to Mansanet-Bataller and Pardo (2008) and the International Emission Trading Association (IETA) annual reports (2005-2009).
[19] This has been defined as a reduction by at least 5 percent of the emissions in 1990. For this purpose, Carbon dioxide (CO2) is the reference gas, against which Methane (CH4), Nitrous Oxide (N2O), Hydrofluorocarbons (HFCs), Perfluorocarbons (PFCs) and Sulphur hexafluoride (SF6) are measured. The largest emitters worldwide, the USA, which accounts for 30 percent (IETA, 2007, p. 201) of total emissions, China, which accounts for 8 percent, and India, which accounts for 5 percent, have yet to sign. For updated information please refer to http://unfccc.int/kyoto_protocol/items/2830.php.

emission-reduction targets. Each flexibility mechanism provides a different “unit” of “emission allowance”, which can be used for compliance. These mechanisms are; the Joint Implementation mechanism (JI), under art.6, the Clean Development Mechanism (CDM), under art.12, and the Emissions Trading System (ETS), under art.17.

More specifically, JI provides an incentive to Annex I countries to develop a “green” project in another Annex I country.[20] This project creates an allowance unit, referred to as an “Emission Reduction Unit” (ERU). ERUs can be used for compliance by the first Annex I country. This provides an incentive to both countries to develop industrial activity in an environmentally friendly fashion. Along the same lines, CDM refers to the implementation of “green” projects in developing countries. The associated unit is the “Certified Emission Reduction” (CER). CERs are issued for projects that have to been approved by the Executive Committee of the “CDM Board for Projects”, which is the institution that issues these allowance units. The Kyoto Protocol defines the role of this mechanism: *“to assist Parties not included in Annex I in achieving sustainable development and in contributing to the ultimate objective of the Convention and to assist Parties included in Annex I in achieving compliance”* (please refer to the full text, which is available at http://unfccc.int/kyoto_protocol/items/2830.php).[21] Another important compliance unit is the “Assigned Amount Unit” (AAU). AAUs are allocated by national governments to regulated installations, according to their fixed target.[22] According to the third flexibility mechanism, namely the Emissions Trading System (ETS), when a company has a surplus of allowances of any type of compliance units, it can sell it, in an organized market, to another company, who may need it to cover a deficit in its emissions-allowances balance.[23]

[20] Countries that have ratified the treaty can be divided into three categories. First, Annex I countries are the industrialized economies and the economies in transition. They need to cover their emissions costs. Annex II (subcategory of Annex I) includes the advanced economies of OECD, excluding those that were economies in transition in 1992. They need to cover the emission costs of the third group, which includes the developing countries. These countries are not required to reduce emission levels because that would slow down their development, since industrial capacity is connected to emissions levels.

[21] These countries might not be regulated under the Kyoto Protocol, but they play a crucial role in the overall emissions reduction, mainly due to CDM and Carbon leaking (Lecocq and Ambrosi, 2007). Regulated companies can either develop “green” projects in developing countries, which is beneficial for both in the sense that the first country is issued CER units while enhancing the industrial activity in the developing country, or they can transfer their industrial activity to places, where they do not need to provide compliance units. In addition, they mention that the projects that are the most beneficial in environmental sense, are not necessarily projects that allow for higher growth in developing countries.

[22] This refers to the companies related to the following activities: combustion plants, oil refineries, coke ovens, iron and steel plants and factories making cement, glass, lime, brick, ceramics, pulp and paper.

[23] Another unit that can be used for compliance is the “Removal Unit” (RMU), often referred to as “sinks”. They are related to “land use” and cannot be traded. Along the same lines, “Verified Emission Reductions” (VERs) are units that cannot be used for compliance but they can be traded in voluntary

It follows that each country needs to provide a well-specified plan for regulating its emissions, and each year's emissions need to match the associated allowances. After establishing a registry, where emissions and allowance units are registered, countries are eligible to start trading. The Independent Transaction Log (ITL) monitors the trading process. At the end of each period, governments need to surrender, and consequently cancel, the allowances that correspond to their emission levels. In case no sufficient number of allowances, *R*, can be provided to match the level of emissions, *E*, a penalty, *P*, is payable for each excessive ton of CO_2.[24] R is the balance of allocated allowances, *AAU*s, compliance units due to development of "green" projects, *ERU*s and *CER*s, and land-based units, *RMU*s. Allowances can also be traded, where the net balance between purchases, *P*, and sells, *S*, is taken into account, along with any stored compliance units, *B*. Consequently, each country's inventory can be written as:

$$R = AAU + ERU + CER + P - S + RMU + B = \begin{cases} \geq E & \Rightarrow \quad Commitment \\ < E & \Rightarrow \quad Penalty = P * (E - R) \end{cases}$$

When the number of allowances provided equals or exceeds the annual emissions, the country is in line with its commitment and no further action is required. When *R* is lower than *E*, the compliance units provided do not match the annual emissions and then the country needs to pay a penalty. This penalty does not "release the company from the obligation" to provide the allowances. Therefore, the country needs to purchase the remaining number of allowances from the spot market, while a penalty for each missing unit is still payable. [25]

1.4.2 The European Union Emissions Trading System (EU ETS)

Description

In Europe, the main effort to tackle the emissions problem embraces the third flexibility mechanism, the ETS. For this purpose, the "European Union Greenhouse Gas Emission Trading System" (EU ETS) has been set up, creating a new "commodity" market. In accordance with AAUs, the unit traded in the EU ETS is the "European Union

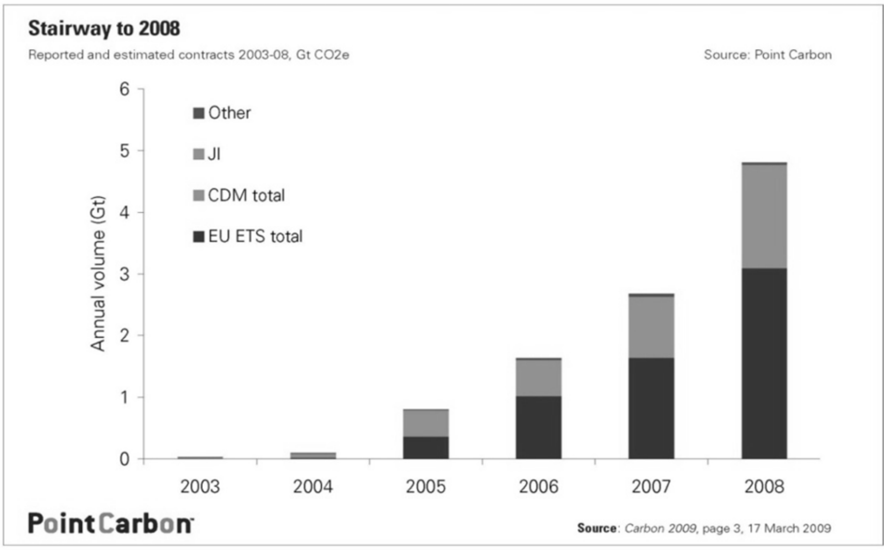
markets by market participants, such as governments and environmental organisations, or other institutions that voluntarily take responsibility of their emissions (see, inter alia, Taiyab, 2006).

[24] The penalty is €40 (€100) per excessive tonne of CO_2 in Phase I (Phase II), which *"...shall not release the company from the obligation to surrender an amount of allowances equal to those excess emissions..."*

[25] Banking, which is defined as storing of allowances for future use, is not allowed between Phase I and Phase II, but it is allowed between Phase II and III, as well as between the years of the commitment period.

Allowances" (EUAs). Purchasing one EUA entitles the holder to emit one ton of CO_2 equivalent GGs. In addition, the phases of the market coincide with the phases of the Kyoto Protocol, underlying the significance of the market, both locally (i.e., in Europe) and worldwide. Gradually, this market has gained complexity and maturity. More sophisticated products have been developed and trading activity has been boosted significantly. In particular, EUAs futures contracts, especially the end of the year contracts, appear to be the most liquid. In fact the EU ETS is the largest CO_2 trading system worldwide (Benz and Hengelbrock, 2008) and the most important of the three flexibility mechanisms (see, inter alia, Gupta et al., 2008; Ellerman and Buchner, 2008; Grubb et al., 2010). Figure 3.1 below presents the relative proportion of the market.

Figure 1.1: Relative Size of the EU ETS



Annual Volumes and Values of Transactions on the Main Allowances Markets

	2005		2006		2007		2008	
	Volume ($MtCO_2e$)	Value (MUS$)	Volume ($MtCO_2e$)	Value (MUS$)	Volume ($MtCO_2e$)	Value (MUS$)	Volume ($MtCO_2e$)	Value (MUS$)
EU ETS	321	7,908	1,101	24,357	2,060	49,065	3,093	91,910
New South Wales	6	59	20	225	25	224	31	183
Chicago Climate Exchange	1	3	10	38	23	72	69	309
Total	**328**	**7,971**	**1,131**	**24,620**	**2,108**	**49,361**	**3,276**	**92,859**

The first panel of Figure 3.1 compares the size of the EU ETS with other ETSs and the other flexibility mechanisms, namely the Joint Implementation, *JI*, the and the Clean Development Mechanism, *CDM*. The metric unit displayed is Gigatons, *Gt*. The second panel of Figure 3.1 reports the relative volumes and values of the three main Carbon Exchanges, namely the EU ETS, the New South Wales and the Chicago Climate Exchange. The units used are Megatons and thousands of US$ for volume and value respectively.

The first panel in Figure 3.1 shows that the EU ETS is the main action globally in achieving emission reduction targets. In 2008 it accounted for around 60 percent of the global action against CO_2 emissions, in terms of volume of compliance units (Mansanet-Battaller and Pardo, 2008). The second panel of Figure 3.1 reports the volumes and values of transactions on the main Carbon markets. The EU ETS exceeds by far the other two both in total volume of transactions and in total value.

In the EU ETS, each country allocates to each compliant company a certain amount of European Union Allowances (EUAs), which is their limit in CO_2 emissions.[26] Before the beginning of each phase, each country needs to make a plan, the National Allocation Plan (NAP), concerning the total number of allocated allowances, as well as the way in which they will be distributed to the regulated installations.[27] NAPs must be submitted for approval to the European Commission, EC, 18 months before the actual beginning of the phase. The assessment and acceptance, or potential correction, period is three months. By the end of March, each year, each member state needs to submit to the EC a report, about the previous year's emissions. The final day for the actual delivery of allowances is the end of April. The following diagram by Mansanet-Bataller and Pardo (2008) shows the process.

Figure 1.2: Compliance process

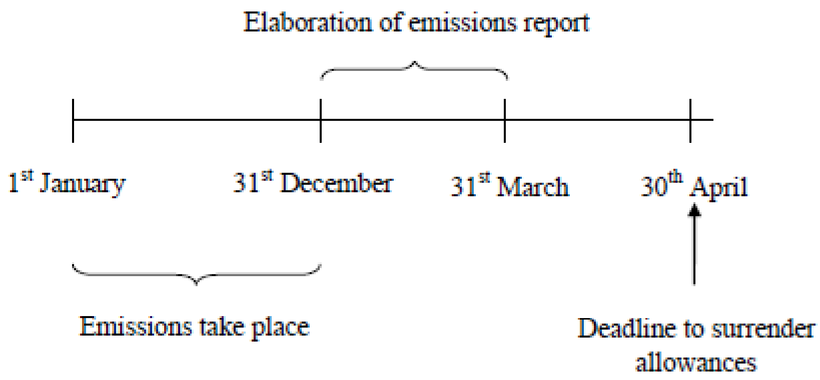


Figure 3.2 shows the sequence of events from when the emissions take place till the delivery of the allowances.

Consequently, the EU ETS is a "cap-and-trade" market, where the overall quantity is predetermined. This results in a price formation process that is heavily influenced by policies. Kruger et al. (2007), commenting on this issue, discusses the centralized-

[26] Several independent studies argue that there has been an over-allocation of Carbon Allowances, with a direct environmental and financial impact, especially in Phase I (e.g., Daskalakis et al., 2006, 2009).
[27] 95 percent of the allowances in Phase I and 90 percent in Phase II are allocated freely, while the remaining proportion is auctioned (Ellerman and Buchner, 2007).

decentralized character of the market. According to their study, the EU ETS is balancing between a completely decentralized market, in the sense that each country can decide on its own emission-reduction plan, and a wholly centralized market, in which the total quantity for all member states is regulated by EC. The ultimate goal is to make CO_2 emissions expensive, in order to provide individual companies the incentive to develop environmental friendly technologies. Along the same lines, Viswanathan (2010) argues that EUA prices, especially in the futures market, reflect investors' beliefs about future "green" technologies. Consequently, the price of EUA futures contracts should fall when cost effective "green" technologies are expected to be utilized. In the same study, he also argues that the Carbon market needs a regulatory framework that can keep a fine balance between market innovation and liquidity, where the viability of the market is of utmost importance.

Trading System

Trading, in the EU ETS is monitored by the Community Independent Transaction Log (CITL), which, similar to previous experiences, such as the Chicago Climate Exchange (CCX), is organized into accounts. Each member state's registry is connected to its account, where any imbalances are monitored.[28] Compliance monitoring is achieved through these accounts and refers only to regulated sectors. However, any physical person is allowed to enter the market and start trading. In either case, allowances that cannot be stored due to banking restrictions, need to be cancelled. According to 2003/87/EC (art. 13), *"each member state must cancel the allowances that are no longer valid and that have not been surrendered and cancelled"*.

Furthermore, Carbon trading in Europe takes place either OTC or in organized markets. More precisely, the OTC market pre-existed and acted as a reference point for the organized exchanges (e.g., Daskalakis et al., 2006; Ellerman and Buchner, 2008). Even in the early stages of the market, in Phase I, OTC transactions were the prevailing determinant of price formation (e.g., Cappor and Ambrosi, 2006, 2007, 2008, 2009; Convery et al., 2008; Alberola et al., 2008) especially in the Futures contracts market (Convery and Redmond, 2007), in the voluntary markets (Hamilton et al., 2007, 2008) and under banking restrictions (Alberola and Chevallier, 2009). The presence of OTC transactions was so profound that their price indices acted as a reference point in Carbon

[28] Since April 2009, all registries in the EU are linked to the United Nations Carbon market, under the Integrated Trading Log (ITL).

pricing. In addition, some markets allow OTC EUA holders to register their positions, by entering the organized trading process.[29]

As the EU ETS increases in size and gains complexity, the number of operating organized markets should be able to accommodate the existing market forces. Consequently the importance of OTC transactions is expected to moderate. According to the EU ETS market regulation there is no restriction on how many markets or trading platforms should exist at any particular moment in time. This emphasizes Kruger et al.'s (2007) comments about the centralized-decentralized character of the market, from a different perspective. Although the Carbon market is a united, common, market in terms of the asset being traded, trading is distributed over various places. Each organized market accommodates regional needs, but since the same asset is being traded and its total quantity is restricted, the total emissions in Europe are monitored centrally by the EC.

The asset being traded in the EU ETS is the EUA, which is common in all markets. This is the EUA spot contracts and the delivery is physical, between accounts. The unit of and the minimum price tick (€0.01) are the same in every exchange, but the size of the contracts, defined as the number of EUAs, might differ across trading platforms. In addition, other contracts, such as futures, forwards and options, which can be traded in the organized market, differ slightly, mainly on stylized facts, such as the contract size or the maturity date and, consequently, the life of the contract. EUA spot contracts can be traded in various exchanges, such as in BlueNext (Paris), in the Energy Exchange of Austria (EXAA, Vienna), in Nord Pool (Norway), in the European Energy Exchange (EEX, Leipzig), in the European Climate Exchange (ECX, London) and in Gestore Mercato Elettrico (GME, Rome). Futures can be traded in ECX, Nord Pool, EEX and BlueNext. Options started being traded in ECX on 13 October 2006 and since then, more markets have incorporated them into their trading platforms. Since Phase II, all compliance units can be traded in most exchanges, while some markets offer financial products of higher complexity, such as swaps and futures contracts on the spread between EUA and CER. Consequently, potential overlapping periods could transmit both information and risk, and markets might have a shared price impact (see Benz and Hengelbrock, 2008).

[29] Some of the most well-known early indices include the *European Carbon Index* and of the European Energy Exchange (EEX), as well as the various indices of the London Energy Broker's Association (LEBA), which is an association of 10 members, providing information for all energy commodities.

Figure 3.3, below, shows that futures contracts constitute the vast majority of trading, 77 percent, especially in Phase II, where the relative volume is twice larger than in Phase I. In addition, a small proportion, 4 percent, of trading volume relates to options on EUAs futures. OTC transactions account for a considerably large market share, 19 percent. In addition, the main trading activity of the EU ETS is concentrated on ECX, which accounts for 95 percent of all contracts being traded. Nord Pool and EEX follow with 3 and 2 percent market shares, respectively.

Figure 1.3: Relative volumes of Spot and Future Contracts

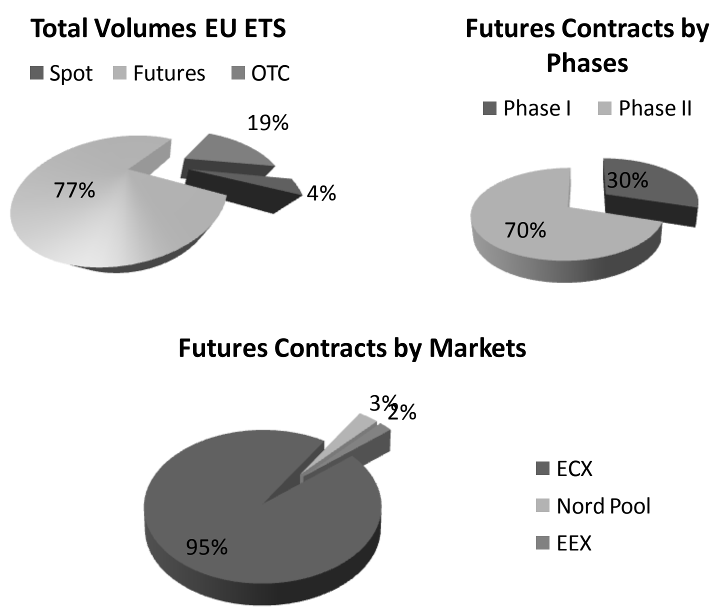


Figure 3.3 presents the relative volumes (in tons of CO_2) of spot and futures contracts in each phase and across markets. Source: Mansanet-Bataller and Pardo (2008).

An important characteristic of the market is that block trades are allowed. These trades offer the opportunity for market participants to bilaterally negotiate trades, without revealing any information to the market and therefore cause an adverse price revision, as long as the order exceeds a minimum threshold.

Another important aspect that is particularly relevant to the present study is that OTC transactions can be registered with the organized market to mitigate counterparty risk. In Nord Pool they can enter directly, taking part in the normal trading process. However, in ECX they need to use either the Exchange for Physical (EFP) or the Exchange for Swap (EFS) facilities. EFP provides the opportunity, after a mutual agreement, for OTC transactions to be cleared by the London Clearing House Clearnet

(LCH Clearnet). The new contract created is standardized and the delivery is physical. Therefore, EFP is mainly used to clear futures and forward contracts. EFS provides the opportunity to register a non-physical OTC position with the organized market, without the physical delivery of the underlying asset. Therefore, it is mainly used to clear financial contracts, such as options and swaps.

1.5 Present Analysis

Summarizing, the EU ETS has some features that distinguish it from other financial markets. First, it appears to be rather illiquid, at least in Phase I, compared to other well-established markets of other commodities or other financial products. Second, it is a “cap and trade” market and refers only to specific energy-demanding, industrial sectors. Consequently, the overall quantity is politically influenced, which is expected to have a significant impact on prices and trading activity. Third, unlike other financial assets or commodities, EUAs futures contracts have a very long time to maturity and their examination might reveal different economic patterns. Fourth, standardized contracts are being traded in simultaneous, non synchronous, but overlapping trading mechanisms. Fifth, OTC transactions are allowed to register with the organized market. This allows for information flow, which is expected to affect trading patterns and pricing. Consequently, econometric modelling that might be successful in other markets might not hold in the EU ETS without further empirical adjustments.

These characteristics, along with the market’s economic importance in determining energy and emission prices, constitute a unique trading environment that has not been investigated adequately, in terms of intraday market dynamics. It appears that there is no relevant study that models duration. This is the main concern of Chapter 4, where the ACD framework is empirically extended to account for market stylized facts. Potentially different trading patterns from other financial markets are investigated. Market liquidity increases exponentially (Daskalakis et al., 2009) and trading intensity is employed to account for non-linear effects of market activity on duration. A new model that focuses on the potential asymmetric effects of past durations is proposed, where the size of the impact of past realized arrival times is allowed to vary across different regimes of trading intensity. Thus, an exogenous variable, such as trading size, is allowed to indirectly determine duration expectations. The role of OTC transactions is further examined by allowing them to affect the impact of past durations differently from normal trades.

According to the analysis in this chapter, in a unique market environment, such as the EU ETS, further empirical adjustments are needed in duration modelling. This is consistent with Bauwens et al. (2004), who argue that the conditional mean specification contributes to model performance more than the distributional assumptions. Although the flexibility of using a more generic density function increases model in-sample or out-of-sample accuracy, linear specifications, which include piecewise linear specifications, are found to clearly outperform their non-linear counterparts. In-sample goodness-of-fit or out-of-sample forecasting accuracy is improved even further when OTC transactions are considered, especially in Nord Pool and in Phase II. In contrast, ECX appears to be affected more by order flow imbalances, which indicates higher complexity and a more mature trading environment. Moreover, OTC transactions appear to increase the expected duration, either because they carry information and deter other traders from transacting, or because they consume the current levels of liquidity. Consequently some time might be needed for the market to reach a new equilibrium.

These findings are particularly relevant to regulatory authorities, since a better understanding of intraday trading activity can enhance market regulation by reaching a balance between market innovation and liquidity needs. According to Viswanathan (2010), this would provide the foundations for more accurate pricing and would help the market serve its purpose of reducing emissions. A more precise duration model could also improve market making practices in EU ETS. Market makers can better manage risk if they know how long (i.e., expected duration) they are going to be exposed to it, and this would result in narrower spreads. Limit order traders can also benefit from a better duration model, since they can develop better trading strategies by managing the time dimension of “uncertainty of execution”.

In Chapter 5, focus shifts to the informational content of trading intensity and to whether ACD models can be used to identify different types of trades and, consequently, different types of traders. A key idea is the role of time in information resolution, and more precisely when and for how long a piece of information is exploitable. Unlike earlier literature, the proposed model focuses on the process of information resolution rather than on its content. Market is seen as being efficient, in the sense that rationally incorporates information, but imperfect, in the sense that it takes some time from the moment that information hits the market until the moment prices reflect that information.[30] This period of "price adjustment" provides the opportunity to some traders to act before everyone else and make a profit. Consequently, the information benefit is now translated into a "timing" benefit. Modelling such time dimension of information allows for identifying informed trades/traders in a natural way, based on past observable information.

This is the main contribution to the literature, which is particularly important and relevant to various aspects of trading practice. First, it highlights the benefit of possessing information before it is fully incorporated into prices. On the microstructure level, market imperfections are exploitable, even for a very short period of time, while information is obtained at a significant cost. This new "time dimension" of information increases the value of "real time" information, since it can increase profitability of intraday trading strategies. Consequently, this can affect the attitude of market participants towards the timing and the cost of acquiring information. Second, market participants can now identify informed trades by simply observing past transactions. They can extract price relevant information can act upon it, improving the profitability of their strategies. Third, the proposed model can be used for monitoring purposes by regulatory authorities. By identifying informed trading, further action can be taken to protect the market from "manipulation". This can also be applied in real time to adjust the balance between market innovation and liquidity.

Drawing on De Luca and Zuccolotto (2006) and on Hujer and Vuletic (2007), a Smooth-Transition-Mixture of Distributions ACD (STM-ACD) model of duration is proposed. Teräsvira's smooth transition framework is applied to the shape parameter of a Weibull distribution, allowing it to vary across three different regimes of an economically relevant variable. This determines the shape of the distribution (the Weibull nests the Exponential) and consequently the shape of the hazard function. The three recognized regimes correspond to three different types of trades/traders; uninformed, fundamental and informed. The smooth transition function accounts for the learning process between regimes and therefore for hybrid trades. Following Hujer and Vuletic (2007), the shape of the hazard function is then related to a particular type of trades/traders, depending on its slope. However, unlike earlier work, trading intensity, which is observable, is used as threshold variable, and this allows the identification of the regime to which each transaction belongs. Further, focusing on informed trading

[30] Public information is assumed to be incorporated much faster than private information, yet still prices are not immediately adjusted. But first, private information needs to be extracted first, after observing trading history. By definition, this prolongs information resolution.

activity, the probability of the next transaction to be informed can be computed following forecasting the level of future trading intensity.

The most important finding of Chapter 5 involves the role of trading intensity in transmitting information signals. First, trading intensity appears to be sufficient in identifying three distinct regimes. Then, according to the shape of the hazard function, the STM-ACD confirms the theoretical propositions of Easley and O'Hara (1992) and Dufour and Engle (2000b) that increased trading activity is associated with information. According to the shape of the hazard function, low trading intensity, in the form of either low trading size or long duration, is related to uninformed trades/traders, while high trading intensity is related to more informed trading. The higher the trading intensity, the more informed the trade. Consequently, large or fast transactions are interpreted as closely related to information and are, therefore, expected to have a greater price impact. Similarly, based on the findings of Pascual et al. (2004), trading intensity increases following informed trades, but this happens because of higher trading frequency rather than higher trading size. In addition, empirical findings provide further evidence that longer durations are associated with no news, according to Easley and O'Hara's (1992) propositions, and contrary to Diamond and Verrecchia (1987), who argue that longer durations are associated with absence of new information. Findings also support Kyle (1985) indicating that informed traders act strategically by segmenting their trades.[31]

In Chapter 6 the dual pricing impact of trading intensity is further investigated, to account for its information and liquidity content. A new dynamic expectations structural model for intraday price changes is proposed, extending the existing literature in the following ways. First, it deviates from a static approach of modelling intraday prices and Bid-Ask spread, in the sense that it allows price components to vary according to trading activity. The information and liquidity components are revised after trading intensity fluctuations. Second, price revisions depend on the, Bayesian, learning process of market participants, who condition their quoted prices upon expectations about the price impact of future transactions. Consequently, price and spread components are modelled as continuous latent variables that depend on past trading history, public

[31] The average duration of Buys is everywhere found to be longer than that of Sells. Considering that informed traders buy (sell) upon good (bad) news, that means that it takes longer to transact upon the arrival of good news. However, a closer inspection of the data reveals that the Carbon market is a buyer market and therefore, it should be easier for a Seller to find a Buyer than the opposite. Hence, the longer durations.

information and market participants' learning ability. Third, the dynamic character of the model provides a framework to measure the information and liquidity pricing impact of a trade, in a way that the profitability of a market or limit order strategy can be maximized. These components are measurable and are revised after every transaction. Investors can formulate expectations and then take an appropriate market position. Finally, this model further investigates the intraday price formation of Carbon allowances, recognising the pricing impact of market participants' behaviour. This can have numerous implications for regulation and trading in the EU ETS.

In more detail, the model proposed in this chapter extends the original model of Madhavan et al. (1997) and draws on the studies of Grammig et al. (2007), Angelidis and Benos (2009) and Ben Sita (2010). Market makers, in the context of a hybrid market, continuously observe trading history through trading intensity, trying to extract price-relevant information. They quote their prices based on liquidity considerations and the post-trade effect of their actions. Similar to Madhavan et al. (2007), intraday returns are driven by both an information and a liquidity component. In this thesis, however, these are allowed to be determined by dynamic expectations of trading intensity, and by the expected exposure to risk. The employed risk measures account for both level of risk and time of exposure. Thus, the model can be used to define the components of estimated spread and intraday volatility, as well as to help identify the most appropriate type of order, between Market and Limit Orders.

One of the main findings is that spread components follow intraday patterns that seem to be adequately explained by the dual role of trading intensity. Trading intensity is positively related to the information component, and, probably due to economies of scale or the illiquid character of the market, expectations of higher trading intensity have a decreasing effect on price changes. The higher the expectation, the higher the price impact of a trade, which confirms the empirical findings of Dufour and Engle (2000b). In addition, and consistent with the literature, risk, in the form of price volatility, seems to affect only the liquidity component of the spread and not the information component. Another main finding is the confirmation of the positive relation between limit orders and spread width. However, it appears that a limit order strategy could be profitable only when trading intensity is low. Otherwise, the information component indicates that the midpoint deviation would be larger than the spread. Finally, Buy orders, especially the large ones, are associated with a higher adverse-selection component and wider spreads than Sell orders.

These findings are particularly relevant to various aspects of trading process in the EU ETS. First, trading practices can be improved, as investors could measure more precisely the price impact of their trades, which might vary over time. They can account for behavioural aspects of trading in real time and they can formulate trading strategies that take into account the actions of other investors. This practice can also be beneficial to investors who possess price unresolved information, as they can adjust their trading to current liquidity levels and sensitivity towards new information, thus maximizing their profits. Second, market making can be further improved, as this model can indicate when market orders can be profitable. Dealers can adjust spread width according to liquidity and information, thus becoming more efficient in managing both. This, in turn, would result in narrower spreads. Finally, a better understanding of trading practices and their pricing impact can improve regulation and monitoring, and thus the market can be more efficient in achieving its goal of reducing emissions. This model proposes a natural measure of market sensitivity towards information and liquidity, which are both highlighted in the Carbon market literature. Regulatory authorities can develop policies that can manage both and allow the market to reach a balance between “market innovation” and liquidity.

The remainder of this thesis is organized as follows. Chapter 2 provides a critical review of the literature, while Chapter 3 presents the data employed, along with a brief non-parametric analysis. The following three chapters, Chapter 4, Chapter 5 and Chapter 6, present the models employed and the empirical findings. Finally, Chapter 7 presents the conclusions of this thesis.

Chapter 2

Literature Review

2. Literature Review

2.1 Review of Previous Studies

Technological advancement and market modernization have provided researchers and practitioners with data sets of increased frequency. Unlike daily observations, which usually aggregate relative quantities, Ultra-High-Frequency (UHF) data consist of transaction level information, where all transactions are recorded. This gives the opportunity for further insight into the trading process and the microstructure of the market, which has become a major and rigorous field of study. Unlike previous studies, prices, which are generally considered to summarize market conditions in one figure, are no longer seen as macroeconomic phenomena. Instead, intraday dynamics of several market relevant variables, called marks, as well as their intraday seasonal patterns are realized to have a transitory or permanent price impact. On that level, trades are recognized to carry price-relevant information and traders learn by observing past transactions. This way, they formulate expectations that influence their price setting. Consequently, price movements depend, not only on new incoming information but on market participants' strategies as well.

Intraday market dynamics have been extensively examined in sophisticated, developed and liquid markets, but developing or illiquid markets still remain unexplored. Limited studies investigate intraday phenomena in emerging markets, such as the Mexican Derivatives Market (Ben Sita, 2010) or the Athens Stock Exchange (Angelidis and Benos, 2009), and report that these effects are magnified due to liquidity or information shocks. Intraday market characteristics are found to play a significant role in price formation, and market dynamics seem to determine prices along with information about the "fundamental" value of the underlying asset.

One of the most recent, and less examined, markets is the European Carbon market. The European Emissions Trading System (the EU ETS), which is the European part of the global Carbon market, is a fairly new market, which has been developed significantly over the last years and has gained complexity and maturity, as well as sufficient liquidity and market depth.[32] The literature on the microstructure of the market is rather limited and focuses on the intraday price dynamics, employing, though, models that

[32] A short description and further particulars has been presented in the previous chapter, but this study does not have the pretension to provide an in depth analysis of the market. Consequently, the literature references will be limited to present the most relative studies.

have been used on lower frequency data sets and in well-developed markets. The majority of these models reports that further empirical adjustment is needed in order to model the market's stylized facts. Viswanathan (2010) further comments on it, supporting that the Carbon market is sensitive to information and liquidity and that a better understanding of market dynamics would improve market efficiency.

2.1.1 Carbon Market-Early Literature

Early literature in the Carbon market focuses on the description of the market mechanism and the general framework of the Emissions Trading System. Kruger et al. (2007) comment on the centralized and decentralized character of the market, emphasizing the need of a consolidated legal framework.[33] Convery and Redmond (2007) discuss the institutional framework of the market, as well as common facts with other cap and trade markets and the OTC predecessor of the EU ETS. Mansanet-Bataller and Pardo (2008) emphasize the driving forces and the history of the EU ETS. They also provide a short description of OTC and the voluntary Carbon markets. Along the same lines, Chevallier (2010) analyzes the nature and origins of CERs (i.e., Certified Emission Reduction units), as well as their regulatory framework. Furthermore, Burtraw et al. (2002), Böhringer and Lange (2005), Kosobud et al. (2005) and Schleich et al. (2006) develop simulation studies to examine the potential effects of market designing issues on prices and trading.

More recently, Viswanathan (2010) argues that the Carbon market needs a regulatory approach that would strictly regulate the market in order to restrict manipulation, but simultaneously allow exchanges' and market participants' innovations to sustain liquidity. A non-regulated, non-transparent market would be liquid, but inaccurate in terms of price. In contrast, a strictly regulated environment would increase price accuracy, but not liquidity. Both would result in a divergence from the EU ETS's initial purpose. This is based on a seminal study of Kyle and Viswanathan (2008), who postulate that the organized markets improve the aggregation of diverse information, which, according to Hayek (1945), leads to an increased "market efficiency". However, that does not necessarily mean that it increases the "price accuracy" as well. They explain that a highly manipulated market, in terms of increased presence of informed

[33] The decentralized character of the market refers to the fact that each country prepares individually a NAP of allocating allowances among the companies that are regulated. However, since the EU ETS is a cap and trade market these NAPs and the associated quantity of allowances is controlled by the European Committee, which is referred to as the centralized feature of the market.

market participants, might be highly efficient, in terms of incorporating manipulation into price, but the asset is not necessarily priced accurately. In the case of the Carbon market, over- or under-pricing would lead to a highly speculative environment, which would undermine efforts to reduce emissions.

Despite its size and importance, the Carbon market literature has only recently attracted academic attention. Some studies discuss the very nature of the asset and develop models that account for both spot and futures prices. Kosobud et al. (2002) and Daskalakis et al. (2009) argue that EUA is a commodity and should be treated as such. Other studies, such as Uhrig-Homburg and Wagner (2007), Truck et al. (2007) and Daskalakis and Markelos (2008), develop cost of carry models to investigate the price formation process. Uhrig-Homburg and Wagner (2007) report evidence in favour of the cost of carry approach for Phase I, while Truck et al. (2007) show that inter-phase futures exhibit significant convenience yields. Daskalakis et al. (2009), however, argue that the standard framework cannot entirely capture market dynamics and, instead, they propose a two-factor equilibrium model, based on a jump-diffusion process. They argue that it is the best approximation to model intra- and inter-phase options on future prices.

Other studies extend the inter-phase analysis, examining the impact of banking, defined as storing of allowances in one phase to be used in the next, and various other factors of trading activity, such as price development. Various studies (see, inter alia, Rubin, 1996; Schennach, 2000; Godal and Klaasen, 2006; Schleich et al., 2006; Daskalakis and Markelos, 2008; Daskalakis et al., 2009) further discuss and empirically examine the banking restriction between Phase I and II and its potential impact on prices, volumes and trading activity. The main finding of these studies, although each presents results analogous to the methodology employed, is that banking restrictions reduce market efficiency and increase risk levels. Vinocur (2009) re-examines the issue, using intra-day data, and emphasizes the importance of repealing the restrictions.

Along the same lines, several studies (e.g., Christiansen and Arvanitakis, 2005; Bunn and Fezzi, 2007; Mansanet-Bataller et al., 2007; Alberola et al., 2007, 2008, 2009a, 2009b) examine how EUAs interact with commodities. They all conclude that energy commodities' prices are connected to each other and that they tend to follow similar pricing paths.[34] In addition, Carmona et al. (2008) and Kara et al. (2008) show that

[34] In addition, Rezek (1999) and Schennach (2000) examine the impact of technological advancements and electricity demand on the price equilibria of SO_2 prices, providing fairly similar results.

EUAs have an increasingly significant impact, not only on energy commodities, but in all economic activities, since they restrict the availability, or else the price competitive advantage of low-cost energy. In contrast, similar studies on the SO_2-CO_2 relation (see, inter alia, Burtraw et al., 2002; Bohringer and Lange, 2005; Kosobud et al., 2005; Schleich et al., 2006) reject common revisions. Research interest focuses also on the effect of National Allocation Plans (NAPs) on price formation.[35] All studies (e.g., Uhrig-Homburg, 2007; Milunovich and Hoyeux, 2007; Sanin and Violante, 2009; Conrand et al., 2010) that investigate this issue conclude that NAPs have a significant price impact, mainly due to the fact that they determine directly future liquidity.[36] Moreover, the majority of studies (e.g., Mansanet-Bataller and Pardo, 2007; Rotfuss et al., 2009) criticize that, especially in Phase I, NAPs have been really generous and an over-allocation of allowances has decreased their trading value, making the market deviate from its initial aim - to make energy consumption expensive.

The majority of the afore-mentioned studies report that several market dynamics have a significant influence on price formation, which sometimes is more influenced by trading patterns than by macroeconomic information about the "fundamental" value of emission allowances. According to Viswanathan (2010) the impact these trading patterns, might be magnified by the actions of market participants. The market microstructure literature extensively discusses various intraday phenomena and provides the tools for an in depth analysis of intraday dynamics, which are not extensively analyzed in the Carbon market.

2.1.2 Intraday price formation

Empirical market microstructure literature recognizes information and liquidity as the driving forces of intraday price formation. Both are particularly important in the presence of market specialists, especially in a relatively illiquid environment such as the Carbon market (see, inter alia, Benz and Hengelbrock, 2008) where dealers have been introduced to support liquidity. Market's equilibrium is determined by the trading strategies of the market participants, according to their access to relevant information and their liquidity desires. Access to information and marginal levels of liquidity determine dealers' risks and consequently investors' costs mainly in the form of the Bid-Ask spread.

[35] As it will be explained further in Chapter 3, NAPs are plans, according to which, individual countries try to control their emissions. Every year they are submitted for approval by the European Committee, and consequently they are important determinants of the available quantity of allowances in the market.
[36] For further market developments related to NAPs please refer to Zhang and Wei (2010).

Unlike Stigler (1967), who sees trading costs as a sign of market “imperfection”, Demsetz (1968) raises the importance of spreads as a natural outcome of trading activity. Market makers provide a constant demand and supply in the market and they need to be rewarded for bearing the risks and the costs of immediate execution of trading. This is the price of “immediacy”, which implicitly connects costs with trading activity. Therefore, apart from explicit trading costs, such as market charges, the time dimension and the size of a transaction are seen for the first time as inextricable determinants of quoted prices and, consequently, of costs. Therefore, trading costs can no longer be ignored or be seen as mere market frictions since they are directly related to trading activity.

The presence of Bid-Ask spread has a non trivial impact on transaction prices and on the time series properties of asset returns. It may create spurious volatility due to the Bid-Ask spread bounce, which is defined as the movements from the Bid to the Ask depending on whether the transaction is Buyer or Seller initiated, and serial correlation in returns (Roll, 1984). In addition, intraday price setting is affected by current levels of trading activity (inter alia, Garman, 1976; Stoll, 1978) and asymmetric information (inter alia, Glosten and Milgrom, 1985; Kyle, 1985). These are the most widely recognized components of the spread that they are considered to drive the intraday price formation. The investigation of these components appears to be critical for investors, financial authorities and researchers.

In earlier studies the spread was viewed as a consequence of the dealer’s need to recover fixed transaction costs as well as normal profit arising from traders’ desire of immediate executions of orders. These studies were trying to determine empirically which variables can capture cross-sectional variation in spreads and the components of the spread. The first cost identified in the literature is the order-processing cost. This is a fixed cost component and quote action associated costs such as installation costs, exchange seat commissions, labour costs, information provision coats, or clearing commissions can be subsumed under order-processing costs (Landsiedl, 2005). Generally, according to market microstructure theory, the order-processing cost is a compensation cost (labour, equipment costs, etc.). It is a fee charged by the liquidity supplier for standing ready to match Buy and Sell orders.

Demsetz (1968) defines order-processing cost as the sum of the buying premium and the selling concession linked to order execution, and compares Bid-Ask spreads to the inventory mark-up of retailers or wholesalers.[37] Using a sample of 200 securities traded on the NYSE recorded for two days, and a simple model, his analysis strongly indicates that Bid-Ask spreads depend negatively on the intensity of trading activity and positively on the price of the transacted security. More specifically, higher transaction frequency reduces the fixed cost per transaction and the waiting costs, while higher asset prices increase the variable, trading volume dependent, cost linearly. The basic assumption is that the per dollar trading cost is constant.

Demsetz (1968) first recognizes trading costs as the cost of market makers for providing "immediacy". Transaction size and frequency are implicitly assumed to be connected with quoted prices, although they are not explicitly modelled. Therefore, costs are considered to be directly proportional to trading activity and can no longer be ignored. This cost component is largely fixed and therefore is expected to decrease with increased trading activity. Huang and Stoll (1997) and Ahn et al. (2002) postulate that the per unit order-processing cost is reciprocal to trading volume. However, diversified market making in a portfolio of assets could amortize more efficiently this cost and weaken this relation. In addition, in a highly competitive market Bid-Ask spreads should equal the expected marginal cost of providing liquidity, in which case order-processing cost may be irrelevant (Bollen et al., 2004). This is consistent with Benz and Hengelbrock (2008) and Frino et al. (2010) who report a small fixed cost component in both ECX and NP, which further decreases as liquidity increases.

2.1.3 Theoretical Models

The second cost identified in the literature is the inventory-holding cost. Inventory-holding cost, can be considered as of market makers' compensation for taking undesired inventory positions. It is the risk of holding inventory stock. The associated studies, referred to as inventory models, investigate the impact of inventory-holding cost on the spread without ignoring the order-processing cost. In these models, the trading process is a matching problem in which the market maker, facing a risk of incoming order flow imbalances, uses the price to balance supply and demand over time. The key factors are the inventory position and the uncertainty about incoming order flow. Market makers achieve inventory control by shifting the quotes to elicit the imbalance of Buy and Sell

[37] Certain securities are available for purchase by retail investors from dealers who sell the securities directly from their own accounts. The dealer's only compensation for the sale comes in the form of the mark-up, the difference between the price the security was purchased at and the price the dealer charges to the retail investor. The dealer assumes some risk by acting in this capacity, as the market price of the security in his or her inventory could drop before he/she is able to sell it to investors.

orders. The focus is on dealers' optimisation, according to their attitude towards various facts that are discussed below. The risk they face can be either in the form of excessive carrying cost, in the form of excess inventory, or loss of sales, when inventory levels cannot support sharp changes in demand or supply. This particularly relevant in EU ETS where market makers have been introduced to support liquidity. This is a rather illiquid market and their inventories are expected to be significant price determinants.

"Inventory" Models

The first studies dealing with inventory-holding cost are those of Tinic (1972) and Tinic and West (1972, 1974). Tinic (1972), using data of 19 trading days (March 1969) from the same market as Demsetz (1968), examines the impact of liquidity on spreads. He develops a three-way decomposition inventory carrying-cost model.[38] Following Demsetz (1968), he includes competition in his estimations, but instead of the number of exchanges on which a stock is listed, he first introduces the Herfindhal Index (HI) of concentration, in order to describe the distribution of trading activity among market participants. His main findings indicate that spreads increase due to increased price of liquidity, caused by increases of the underlying asset prices, market concentration and decrease of diversification of risk. Moreover, Tinic and West (1972, 1974) provide further empirical evidence concerning the Bid-ask spread components. In their first paper (Tinic and West, 1972) they employ data from NYSE for six days, while in the second (Tinic and West, 1974) they examine Toronto Stock Exchange (TSO). They aim at providing additional evidence concerning the impact of direct inter-dealer competition on the price. In their findings they postulate that price, trading activity and intensity of competition are the basic determinants of the spread.

Extending on these ideas, two separate frameworks have been developed to analyze the Bid-Ask spread components, emphasizing the inventory-holding cost. The first one is the single dealer framework, in which the market maker/dealer has monopolistic power on trading activities and determines the spread (Amihud and Mendelson, 1980). In the second one a competitive multi-dealer (Ho and Stoll, 1983) framework, extended on a multi-period (Ho and Stoll, 1981) setting is examined, where a dealer is considered as a

[38] In this early stage of microstructure literature, prices are seen to be heavily affected by dealer's optimization problem concerning their open inventory positions. Therefore, Tinic (1972) decomposes inventory price impact into the following three components: i) Factors affecting the cost of positioning, such as price, trading volume, institutional activity. ii) Factors affecting the cost structure of the entire unit, such as capitalization, size, nature of specialty portfolio. iii) Factors influencing the profit margin, such as exchange surveillance, indirect competition, inter-market rivalry.

market participant who faces competitive pressure in quoting his/her prices. In either case, market maker/dealer seeks to achieve the optimal inventory level by balancing the inventory-carrying costs against the opportunity costs of lost sales.

In the monopolist setting of Amihud and Mendelson (1980), arrivals of Buy and Sell orders are assumed to arrive randomly, following an independent Poisson processes, while the quoted Bid and Ask prices are shown to be dependent on the dealer's stock inventory. They conjecture that the quoted prices are monotonically decreasing functions of the inventory in hand.[39] Previous studies, such as Demsetz (1968), recognize the stochastic character of the arrival of orders and its importance in "temporal" market microstructure. However, their analytical focus on the trading needs of the market participants prevents them from explicitly modelling it.[40] Garman (1976) first incorporates these ideas in a one period model, where there is only a single market maker who receives, executes and clears all orders, trying to maximize his/her profit per unit of time. Simultaneously, the market maker tries to avoid market failure, which can occur by either no cash or no inventory. The arrival time of Buys and Sells is asynchronous and order flow imbalances, which can occur at any point in time, are the main source of risk. However, according to Garman's (1976) model, the dealer can only set prices once, in the beginning of the trading process, and, therefore, his/her main concern is to "be prepared" for any "midstream" excessive variations in incoming orders. Therefore, formulation of expectations of the arrival of these orders is of utmost importance in order for the dealer to "survive".

Garman's (1976) model, although it is innovative in the sense that it emphasizes the price impact of order flow variations and recognizes them as independent stochastic processes, it involves several strong assumptions. A very strong one is that dealers cannot borrow cash and consequently they can run out of it, which is a market failure. Amihud and Mendelson (1980) extend Garman's (1976) model by relaxing that assumption. A dealer is assumed to set exogenously an upper and lower bound to his/her inventory position and market failure, simply, cannot occur. Therefore, his/her

---

[39] Stochastic processes are used for modelling random events in time that occur to a large extend independently of one another. It is one of the most important models used in queuing theory. For further information see Ross, S.M. (1996) Stochastic Processes, 2nd ed, New York: Wiley p.59

[40] The term "temporal" microstructure refers to Garman's (1976) propositions that temporal imbalances in incoming order flow might deviate market makers significantly from their desired inventory positions. Then, they will change their quotes in order to reverse that effect and they will return to normal prices after equilibrium is restored. This results in a short-term price variation due to inventory positions and therefore the effect is only temporal. As it will be discussed below Hasbrouck (1988, 1991a, 1991b) confirms this idea.

incentive is shifted from "market failure" to profit making.[41] The optimal strategy for the dealer suggests a preferred inventory position, supposing that the dealer can define it, based on the available information. If the dealer aims at liquidating his/her position he/she could increase (decrease) the Ask (Bid) price in order to increase (decrease) investors' demand (supply). Order flow imbalances are still the main source of uncertainty, the analytical focus now shifts to inventory positions. Consequently, this model suggests that there is a direct bond between incoming trades, inventory positions and quote setting, where spreads increase with inventory positions in an analogous fashion (Madhavan and Smidt, 1993; Hasbrouck and Sofianos, 1993).

However, a monopolistic framework imposes two unrealistic restrictions. First, optimal position and consequently price quoting depends only on the volume of incoming orders in one or the other side of the spread. Dealers need to compensate for these imbalances and therefore they charge a spread. According to the afore-mentioned models, they would charge the same cost for any given price, as inventory plays a "buffer" role. The optimal position, though, depends on various other factors (e.g., Stoll, 1978) that in the monopolistic framework they are not allowed to have a price impact. Second, these models describe spreads as reflecting the dealer's market power and they account for unit transactions. However, in the presence of competition, spreads would effectively reach zero. This makes spreads a market trading cost.

In contrast, Stoll (1978) places the dealer in a competitive environment. The market maker is seen as another market participant, with a given portfolio, who is willing to deviate from his/her ideal position, to accommodate trading desires of other market participants. Doing that, he/she is exposed to risk, for which he/she needs to be compensated and thus charges the spread. The difference with previous studies is that dealers' behaviour deviates from being risk-neutral and they are considered to be risk averse. This means that they are willing to undertake extra risk for analogous return. Consequently, spread is no longer a sign of market power but reflects dealers' costs for bearing that risk.

This cost is similar to what Demsetz (1968) calls cost of "immediacy", which, according to Stoll (1978), can be decomposed into carrying, order-processing and

[41] Deviations from the dealer's preferred inventory position result in price changes. This happens simply because the dealer wants to return to that preferred position. There is no risk of market failure and therefore the incentive is to maximize profit. This profit making is not a result of speculation, but it is a compensation for inventory deviations due to incoming trades. Therefore spreads reflect the dealer's market power.

asymmetric information costs. Carrying cost measures the risk of exposure. It consists of price change, order flow imbalances, storing and loss of sales costs. Order-processing cost summarizes any fee related to the trading process and it is largely fixed. This means that economies of scale can be developed decreasing the, per unit, cost for larger transactions. In combination with the fact that excessive exposure, in the form of more or less inventory, increases the actual spreads, this model conjectures that there is an optimal level of transaction size that minimizes trading costs. Finally, asymmetric information cost accounts for the fact that the dealer might transact with traders who possess superior information, concerning the “fair” value of the underlying asset. This will definitely result in losses, caused by unfavourable price movements. Dealers can compensate by charging an increased spread to all transactions, in order to balance the loss with excessive gains from transacting with uninformed transactions. This cost component has been developed extensively in the literature and it is particularly relevant to this study. Therefore, it will be further developed in this and the following chapters.

Stoll’s (1978) model, although revolutionary, has a simplicity that raises concerns about its generality. One of the main drawbacks, present in previous inventory models, such as in Garman (1976) and Amihud and Mendelson (1980), is that the time horizon of risk exposure is fixed at one period. This simplifies the decision process, since dealers know in advance how long they need to keep the inventory for. However, this is a strong assumption for real markets, where the time horizon should determine dealer’s risk aversion as well. Ho and Stoll (1981), extending Stoll’s (1978) model to a continuous time framework, examine the Bid-Ask spread as consisting of two components. The first component is the risk-neutral spread and the second is a risk premium that depends on dealer’s risk profile, transaction size and return variance.[42] They demonstrate that the uncertainty in demand is not eliminated by dealers’ strategy and they have to increase both Ask and Bid prices if they want to build inventories, while in the opposite case they need to decrease both. Their model produces three main findings. First, spreads depend on the time horizon, which is arbitrarily set by dealers, since that determines their level of exposure. Second, there is a risk-neutral, monopolistic, spread that dealers add a premium to, analogous to their risk aversion.[43] Finally, probably the most important property of the model is that prices do not depend on inventories. Similar to Stoll (1978), the fair price is exogenously determined and inventory only affects the size

[42] The risk-neutral spread maximises expected profits for a given stochastic demand function

[43] This is particularly relevant to this thesis, where intraday prices are allowed to vary according to dealer’s expectations. Dealers are assumed to formulate expectations concerning the level of risk and how long they are expected to exposed to it. For further information please refer to chapter 6.

of the spread around it.[44] Several other studies, such as Zabel (1981), O'Hara and Oldfield (1986) and Madhavan and Smidt (1993), extend the continuous time framework to an infinite horizon, where prices are set to maximize the current present value of expected future cash flows minus the carrying cost of inventory.

In the multi-dealer framework, the individual dealer recognizes that his/her welfare depends on the actions of other dealers. Ho and Stoll (1983) examine the impact of inventory-holding cost on securities' spreads under a competitive environment using, contrary to Garbad and Silber (1979), a "transaction by transaction" method. In their study dealers are assumed to face stochastic stock returns and transactions of fixed size.[45] The Bid-Ask quotes are shown to depend on the degree to which transactions are correlated across securities at a given period of time and in a given security over several time periods as well as on the anticipated actions of other dealers. Similarly, Biais (1993) introduces the idea of "incomplete" information, concerning other dealers' inventory positions.[46]

Some studies examine other possible determinants of the Bid-Ask spread, accounting for various market stylized facts. Benston and Hagerman (1974), based on a randomly drawn sample of 314 OTC stocks, report that the unsystematic risk and the number of transactions are positively related to the spread. In addition, Copeland and Galai (1983) show that the spread is positively related to price and return variance, but reversely related to market activity. In their analysis, they characterise the cost of supplying quotes as writing a call or a put option to an information-motivated trader.

However, previous models mainly refer to dealer/auction markets and consequently they do not account for limit orders. O'Hara and Oldfield (1986) account for both types of incoming trades, namely market and limit orders, examining the influence of dealers' risk aversion on their pricing policies for securities. Similar to Ho and Stoll (1981), they consider a risk averse market maker, who adds a premium that accounts for risk aversion to a risk neutral, monopolistic, spread, but they only examine a discrete

[44] For example, supposing that the dealer formulates an expectation that determines the fair price to be 50, then spread could be 49-51 (48-52) when dealers have excessive (less) inventory.

[45] This is in opposition to more recent studies, such as Easley and O'Hara (1987), Engle (2000) and Manganelli (2005), where trade volume is examined as a factor influencing the spread rather than being constant.

[46] According to Madhavan (2000), in dealer markets, inventory variations are simply measured by reciprocal order flow variations. However, that is not the case in a hybrid market, where traders can become market makers in one side of the spread (Madhavan et al., 1997). Therefore, in some cases dealers cannot even extract this information.

framework.[47] They show that the spread can be decomposed into a portion for the known limit orders, a risk neutral adjustment for expected market orders and a risk adjustment for uncertainty concerning market orders and inventories. They demonstrate that a risk-averse market maker may set a narrower spread compared to a risk neutral specialist and that the placement and the size of the spread are strongly influenced by the desired inventory position. Their main finding, though, is that inventory affects spreads only when the dealer faces both price and order flow risk.

In contrast, Cohen et al. (1980, 1981) following a similar methodology with Ho and Stoll (1983), examine the price impact of inventories in hybrid markets, such as the European Carbon market. They describe price movements in securities markets, over time, as being partly the result of underlying economic changes and partly a reflection of the impact of indiosyncratic orders that come in from individual investors. Moreover, they explain finite spreads in a continuous price setting, developing the idea of the "gravitational" pool.[48] Order flow imbalances have now a manageable time dimension that needs to be taken into account. Inventory deviations from optimal positions affect prices only when limit order submissions are driven to switch to market orders, and consequently to increase demand or supply, because of limit order trader aggressiveness.[49]

Summarizing, inventory models explain deviations from the "fair" or "efficient" price as resulting from either risk of market failure and the incentive of "survival" (Garman, 1976), or market power and maximizing profit (Amihud and Mendelson, 1980), or dealer's risk aversion and maximizing terminal wealth (Stoll, 1978), or because of the "gravitational pool" and market participants' aggressiveness (Cohen et al., 1980, 1981).[50] However, despite the different explanations, the basic idea of the inventory

[47] Other studies that employ a discrete framework include Zabel (1981) and Bradfield (1979).

[48] Hasbrouck (2007) defines gravitional pool as: *"When the market Ask price, approaching form above, hits the limit order price, the limit order Buyer is pulled to the Ask-switches to a market Buy order"*.

[49] Limit order trader aggressiveness is a key concept in a broader class of equilibrium models (see, inter alia, Chakravarty and Holdern, 1995; Parlour, 1998; Foucault, 1999), where it is a determinant of the order type selection. The present study covers partially this issue, which is further developed in chapter 6. A new model is proposed which could provide an analytical tool for selecting the most profitable type of order.

[50] In contrast to all previous approaches, there are some researchers like George et al (1991) who reject the assumption that inventory-holding cost should be considered as a component of the Bid-Ask spread. Hanousek and Podpiera (2004) relax this assumption and they argue that this component exists but only in the extreme situation of a general trading pressure. They assert that the traditional assertion only holds true where there is only a single market maker. They also argue that the same risk of undertaking unwanted inventory is shared by a larger number of dealers. In addition, the reaction to inventory is weakened by the behaviour of dealers. However, empirical evidence, presented in Madhavan and Smidt (1991, 1993) and Hasbrouck and Sofianos (1993), report preferred, but not necessarily time variant,

based models' approach focuses on the dealer who faces a balancing problem and tries to moderate deviations of the incoming order flow. These deviations though only depend on the behaviour of the market, in the short run. Therefore the dealers' effect on price, especially expressed by their desired inventory position, can only be temporary.[51] According to Hasbrouck (1998, 1991a, 1991b) and Madhavan and Smidt (1991, 1993), a permanent price impact can only be caused by information. The market maker changes the quotes and therefore the price, depending on the trading costs, on the dealers' previous inventory position and on the net demand to the dealer.

"Information" Models

The next spread component identified in the literature, is the asymmetric information cost. Asymmetric information has a permanent price impact and accounts for the fact that different market participants have access to different levels of information. Information models focus on the impact of information on prices, and consequently on spreads, using different methods to proxy the informational content of trades. Generally this cost is considered to arise because of the presence of informed traders in the market. All models consider as informed, traders who possess price-relevant, exploitable, information prior to other market participants. When market makers trade with them they will, on average, incur a loss. Therefore, part of the spread is considered as a compensation for that risk.

This seems to be particularly relevant in the Carbon market, where the relevant literature (Daskalakis et al. (2009) reports information and liquidity shocks, which are expected to be significant price determinants. Dealers in the Carbon market might face a magnified information risk, due to decreased liquidity. On average they will lose money when dealing with better informed agents, but they can usually compensate by trading with other traders who do not posses price relevant information. However, they might be more susceptible when liquidity levels are low.

The asymmetric information models allow for heterogeneously informed traders. If a trade is information motivated, its occurrence will send a signal to the market revealing

---

inventory positions for market makers that have a rather weak, but existing, influence on spread width. Furthermore, Lyons (1993) provides evidence of significant inventory effects in the FOREX market. Along the same lines Madhavan (2000) reports that, although inventory data are difficult to acquire or to extract in hybrid markets, the liquidity effect of past trades cannot be neglected, even if it is not directly related to inventories.

[51] At the end of the adjustment process, price and inventory are completely reverted. Revision is not immediate, but there is no permanent price impact in this model because trades are independent of information.

part of this information. Other traders who observe the market will capture this signal and they will learn from it. This “knowledge” will lead them to formulate expectations and revise his/her opinion about the “fair” value of the asset and therefore prices will be revised accordingly. These price revisions are permanent and they will hold until a new piece of information is revealed[52] Consequently, the trading process is viewed as a “game” involving different types of traders, according to their access to information. Early theory tries to explain the source, the flow and the destination of the information by studying price fluctuations, order flow innovation and other transaction related liquidity parameters.[53] Theory differentiates between three types of agents on asset markets; noise traders, informed traders and market makers.[54]

The basic idea was firstly developed by Bagehot[55]. In his study, he explains why investors as a whole lose from trading and why informed investors win. He argues that market makers will always lose to informed traders with superior knowledge when fulfilling their duties. He emphasizes the time dimension of information and the time it takes for any incoming piece of information to be incorporated into prices, arguing that, even without explicit trading costs, spreads would still exist. Rational market makers recover their losses by charging widened Bid-Ask spreads. This way they can compensate by transacting with uninformed noise traders.

His study triggered a rapid expansion of this idea and early models can be divided into two broad categories. One class of model considers that traders who possess any price-relevant information enter the market sequentially and independently. The key idea here is that they have the incentive to exploit their information advantage at once, without being affected by adverse, subsequent, price changes. These models are known as

[52] According to the efficient market hypothesis, prices summarize all price-relevant information up to that time, discounting future, expected cash flows. Therefore, they express expectations, which are seen as to be influenced by past trading history. Consequently, when dealers believe that they have extracted a piece of information, they believe that it will eventually be incorporated into prices. Therefore, they revise their beliefs about the “fair” value of the asset permanently.

[53] Information models are mainly theoretical models based on “game” theory. Their main aim is to describe market dynamics and determine market equilibria. A key idea in these models is the Rational Expectation Equilibrium (REE), which is the market balance that summarizes rational, Bayesian learning, traders. This branch of literature has been developed in parallel with the empirical microstructure literature. Considering that the aim of this study is mainly to empirically examine various stylized facts of the European Carbon market, the most relevant literature refers to the empirical examination of various issues. However, whenever is considered relevant some references will be discussed. For a detailed discussion of REE literature, please refer to Vives (2008).

[54] Noise traders trade for liquidity reasons in order to rebalance their portfolio, or just randomly, according to their private beliefs (Kyle, 1985). They have common knowledge and information sets containing only public information. (Landsdiel, 2005). Informed traders are risk-neutral “insiders” or have access to private information. (Stigler, 1967). Market makers provide immediacy and trade with some superior knowledge concerning order flow.

[55] The author’s name is Jack Treynor and he uses the pseudonym Bagehot

sequential trade models. They examine the determinants of the Bid-Ask spread in a competitive framework with heterogeneously informed agents. They are characterised by a probabilistic approach, where prices act as signals, according to the semi-strong form efficiency hypothesis, and there is a Bayesian learning problem confronting market participants.[56] The Bid-Ask spread increases with the degree of asymmetric information and decreases as time elapses and the market maker acquires information. In contrast, another class of models, known as strategic, examine the case where a single informed agent can trade at multiple times, trying to cover up his/her actions. This might happen when there are capital constraints or insufficient order book depth. Obviously, these models recognize that trades carry informational content which can be revealed to other traders who observe the market. Therefore, informed traders need to take into account the price impact of their trades. These models maintain that private information provides incentives to act strategically in order to maximise profits. Informed traders choose an appropriate timing and they might prefer to separate their trades in smaller sized transactions, in order not to reveal their private information, which would move the price against them. This strategic behaviour might be affected by the trading mechanism and it may induce patterns in trading activity, returns and volatility (Grossman and Stiglitz, 1980; Kyle, 1985, 1989).

The first sequential model that extends the idea that spreads would exist even without explicit trading costs is the study of Copeland and Galai (1983). They consider the price setting as writing a call or a put option to an informed trader, independently from dealer's inventory position. This way, market makers are seen as trying to maximize their wealth by increasing the gains from liquidity, uninformed, trades, minimizing potential losses by transacting with informed traders. They postulated that spreads increase with price and variance levels, while competition, trading activity and market depth result in narrower spreads. Contrary to inventory models, which consider information is exogenously determined, information models connect trade history with information.[57]

---

[56] The Bayesian school of statistics is based on a different view of what it means to learn from data, in which probability is used to represent uncertainty about the relationship being learned. Prior opinions might influence the interpretation of the data. These opinions can be expressed in a probability distribution over the network weights that define this relationship. For further information please refer to Edwin (1994), Probability Theory: The logic of science. From investors' perspective, the actions of informed traders reveal partially their private information, but the interpretation depends on each uninformed investor.

[57] In inventory models, market makers observe order flow variations and they try to determine expectations concerning potential future imbalances. On the contrary, in information models, past trades

One of the most influential studies is the one of Glosten and Milgrom (1985). They examine the dynamic properties of Bid-Ask spreads and transaction prices in a pure dealer market.[58]. They assume that informed traders have the incentive to exploit their informational advantage at once. The only thing that market makers can do is to estimate a Probability of INformed (PIN) trading and to set their prices accordingly, in order to compensate their losses by trading with uninformed traders. They do so by observing past transactions and their learning is assumed to be Bayesian. They conclude that the Bid-Ask spread quotes differ from the price that would prevail under homogenously distributed information, partially because of adverse-selection. In addition, they assert that the expectations of specialists and of traders tend to converge over time and that superior private information would lead to higher spreads. As in Akerloff (1970), the market might fail in cases of excessive informed trading, because enormously wide spreads would prohibit trading.

Moreover, both studies play an important role in shifting trading interest from the liquidity component of order flow to endogenous information and they emphasize the direction of the incoming trade. However, they still assume non-variant price impact of transactions of different sizes. In contrast, Easley and O'Hara (1987, 1992) differentiate the effect of various transaction sizes. Using transaction data to examine a sequential information model, they provide empirical evidence that informed traders prefer to trade in large amounts at any given price and, more importantly, that transaction size, sequence of trades and trading frequency are significant spread determinants. They observe that larger size transactions are executed at less favourable prices. This might occur because of insufficient liquidity or because they move market makers far from their desired positions (Stoll, 1978; Ho and Stoll, 1981, 1983), or because they are connected to information. In addition, they postulate that volume of trading can be used to identify the type of traders. This idea is particularly relevant to the analysis presented in chapter 5, where it will be developed further.

Sequential models, although they emphasize the informational content of various market related variables, ignore any potential strategic behaviour and the sequence of transactions, which makes convergence to full information problematic. There are many

---

convey information, which can be revealed by order flow variations. Hasbrouck (1988, 1991a, 1991b) emphasize the dual content of trades, while Madhavan and Smidt (1991, 1993) mention the difficulties in distinguishing them.

[58] Specialists perform no brokerage fees and all orders are market orders. They assume that all risk-neutral specialists operate in a competitive environment, in which the expected profits from each transaction are zero and there are no other transaction costs.

informed agents, but they can only transact once. On the contrary, Kyle (1985) is the first to develop and examine a strategic information-based model. He considers a single informed trader in a multi-period framework. Transactions have a significant subsequent price impact that needs to be taken into account in advance. He employs a dynamic model to examine the informational content of prices, the liquidity characteristics of a speculative market and the value of information to an insider. He argues that the informed trader makes positive profits by exploiting his monopoly power optimally in a dynamic context, where noise traders offer a camouflage which conceals his actions from market makers. Prices follow a Brownian motion, assuming the depth of the market to be constant and that all private information is incorporated into pieces by the end of trading. [59] This study draws from Rational Expectation Equilibrium (REE) literature and introduces a strategic pattern in informed trading. It tries to describe how markets incorporate information by combining market makers' price setting with informed strategies.[60]

The same author in 1989 conjectures that noise traders ignore the impact of their trades on prices and given a distribution of private information, prices become less informative when they diverge from competition equilibrium.[61] Along the same lines Jackson (1991) and Rochet and Vila (1994) examine strategic behaviour variations in different microstructure settings and trading mechanisms. Other extensions of Kyle's (1985) model include Back (1992), Holden and Subrahmayam (1992, 1994), Foster and Viswanathan (1996) and Back et al. (2000) who extend the initial model by considering multi-period versions, with multiple competitive traders. In addition, Admati (1985) and Subrahmanyam (1991b) consider a multiple securities framework. They show that diversification in an index mitigates individual asymmetric information, making it

[59] Brownian motion is a mathematical model, a continuous-time stochastic process, used to describe random movements, often called a particle theory. For further information please refer to Karatzas and Shreve (1991), "Brownian Motion and Stochastic Calculus.

[60] Kyle's (1985) model is closely related to REE literature and it can be considered as the meeting point of theoretical and empirical streams of literature. However, the model deviates from the simple setting of Grossman and Stiglitz (1980) in three main points. First, traders are not necessarily described as liquidity traders submitting their orders simultaneously. Second, risk neutrality of uninformed traders is no longer considered to be valid. Otherwise, uninformed market makers would not have an incentive to transact, or to provide immediacy, and therefore the model would collapse. Finally, price formation is explicitly modelled through an auction setting, unlike the REE model which can be considered as an auction model without price formation modelling. For more information please refer to De Jong and Rindi (2009).

[61] This is in line with Dufour and Engle (2000b) who define a liquid market as a market in which information takes longer to be incorporated into prices. According to Kyle (1989) less competition results in less informed trading and therefore prices are less informative. According to Easley and O'Hara (1992) that happens when trading volume is low. High trading volume indicates higher presence of informed traders. For Dufour and Engle (2000b) this is when trades have a lower price impact, which is translated into increased liquidity.

cheaper to trade. Another issue discussed in Foster and Viswanathan (1990) concerns the life of information and the trading strategies that can be developed in the presence of long-lived information.

Several comments can be made concerning the sequential and strategic approaches of Glosten and Milgrom (1985) and Kyle (1985). First, these two models mainly differ in the way in which they treat informed traders. In the sequential model of Glosten and Milgrom (1985), informed traders will trade intensively as if it is their only opportunity to trade and maximize their profits. But, in the strategic model of Kyle (1984), informed traders choose to trade gradually in order not to reveal their private information. Second, a time dimension is implicitly assumed to be present in both models, since prices converge at different rates for uninformed and informed traders. Third, both models capture the same phenomenon; the informational content of trades moves prices, which are assumed to reflect the expected presence of informed traders. Notably, Back and Baruck (2004) combine both models and demonstrate that, under specific time considerations, they can meet when informed (uninformed) traders trade with lower (higher) trading rate. They assert that informed traders might decrease their rate of trading, when they realize that strategic trading can maximize their profits. In addition, Foucault et al. (2003) show that informed traders need to compensate patient traders for providing liquidity, which definitely decreases their incentive to act aggressively.

Another important issue, extensively discussed in subsequent strategic models, is the characteristics of each group of traders. All previous models dissect market participants into two large groups, according to their access to price-relevant information. Informed traders have an exploitable time advantage, whereas uninformed traders possess only public information. However, several studies criticize the naivety assumed for uninformed traders, which are considered to transact only for liquidity reasons, without observing transaction history. They are usually assumed to ignore informational signals in favour of their exogenously determined liquidity reasons. In more detail, Admati and Pfleiderer (1988) relax the assumption that uninformed traders do not manage the time and the size of their transactions. They introduce the idea of discretionary liquidity traders, who, although uninformed, try to extract information from order flow variations and chose accordingly when to trade. This formulates a market setting with three groups of market participants; informed, discretionary and non discretionary liquidity traders. Along the same lines, Subrahmanyam (1991a), who assume risk-averse market makers, and Spiegel and Subrahmanyam (1992), who assume risk-averse hedgers, model uninformed traders as risk averse traders, who request extra compensation for bearing extra risks.

Summarizing, information models recognize that trades might convey information that uninformed traders could extract by observing past trading history. The information they can extract refers to the "fundamental" value of the underlying asset and it changes their expectations concerning the associated future cash flows. Therefore, price revisions due to revision in beliefs are considered to be permanent. This class of model tries to describe how new information is incorporated into prices. Furthermore, combining the information and the liquidity component of trades, described in inventory models, Hasbrouck (1988, 1991a, 1991b) suggests that trades have a dual impact on intraday price formation. The first is associated with liquidity and its effect is transitory and last only till liquidity imbalances soothe, while the second is associated with information and has a permanent price impact, since it revises expectations. In addition, Madhavan and Smidt (1991, 1993), Hasbrouck and Sofianos (1993) and more recently Bowe et al. (2007) and Ben Sita (2010) develop models in which past transactions are allowed to have both a permanent and a temporary impact on prices. They underline the difficulty in distinguishing between them.[62]

2.1.4 "Empirical" Models

Both sequential and strategic price models provide a further insight into the microstructure of the market and the intraday dynamics of price formation. However, their resilience to game theory is their greatest disadvantage. According to O'Hara (1995), *"in order for equilibrium to exist, the dynamics of the "game" played and the "game" itself must be known"*. Their theoretical approach restricts their empirical application and generality. For that reason another part of literature that examines several empirical aspects of market microstructure has emerged.

An excellent benchmark is the "covariance" model, originally developed by Roll (1984). He extends the martingale property of transaction prices to the "efficient", "fair" price, by allowing returns to be affected by the trading process. The transaction price is assumed to consist of a random walk component plus a component that accounts for trading costs, which is covariance stationary (i.e., covariance of order two and higher

[62] The model proposed in chapter 6, resembles to the models of Bowe et al. (2007) and Ben Sita (2010), and aims at describing both the transitory and the permanent price impact of transactions, through a dynamic formulation of expectations. Trades are assumed to reveal private information and act as a mean to formulate expectations for future liquidity levels.

are essentially zero). This is a structural approach that models intraday price movements as occurring by a permanent revision of the efficient price, captured by the random walk component, and a transitory impact of market activity, which summarizes fixed trading costs, as well as any price changes due to liquidity variations. This model is revolutionary in the sense that it recognizes that intraday price variations are governed by both information and trading activity, in a way that it combines revisions in "beliefs" about the "fair" price and market, order flow, "imbalances". It has become popular due to its simplicity and due to the fact that it provides a way to estimate spreads directly from transaction prices. This triggered a rapid expansion of this branch of the empirical microstructure literature. It provides a practical new way to model information and liquidity price impacts directly, in a contemporaneous manner. However, it works under the assumption that dealers face only order-processing cost and therefore it is considered "*a naïve order-processing cost model of the posted spread*" (O'Hara, 1995).

A crucial assumption in Roll's (1984) model is that there is no serial dependence in transaction direction (i.e., Buy or Sell).[63] Choi et al. (1988) relax this assumption and extend Roll's (1984) model by taking into account the serial correlation of trade initiation. Results indicate that their model explains more than 80 percent of the cross-sectional differences in announced Bid-Ask spreads in the Chicago Board Options Exchange. Along the same lines, Bhattacharya (1986), examining a stock portfolio in the NYSE, allows some orders to be executed at the midpoint. Glosten (1987) analyzes serial covariance of price changes assuming that the total spread is the sum of the adverse-selection and gross profit components. He points out that the negative serial covariance of returns arises only from sources other than information. More recently, Stultz (2000) investigates this issue in more detail and maintains that the realized Bid-Ask spread also measures trading costs net of the information component.[64]

Moreover, Stoll's (1989) model is of particular interest, since it gives a way to estimate all three widely recognized spread components. In this sense, the model is "full". Stoll (1989) uses data consisting of transaction prices and price quotations for NASDAQ/NMS stocks. The serial covariance of transaction returns and the serial covariance of quoted returns are modelled as functions of the probability of a price

---

[63] The conditional probability of a trade flow reversal equals 1/2. P{t is a Sell | t-1 was a Buy} = P{t is a Buy | t-1 was a Sell} = 0.5. Flood et al. (1998).

[64] These studies identify potential downward bias in the Roll estimator. However, none provides an adequate explanation of why serial covariance of price changes are so often positive or why the average weekly serial covariance is more negative than the average daily serial covariance.

reversal, employing a two step estimation.[65] In addition, George et al. (1991) show that the spread estimates provided by Roll (1984) and Stoll (1989) are biased due to the fact that they do not account for potentially time varying expected returns. They use daily and weekly data derived from AMEX/NYSE and NASDAQ stocks, respectively, in order to test their hypotheses. [66] In line with Stoll (1989), they employ a two-stage analysis. First, they compute the serial covariances of price changes and second, they regress them on posted spreads. Also, they measure the spread on the basis of serial covariance of the difference between transaction returns and returns at Bid prices, in order to overcome the problem of returns' time variation. Their findings show that order-processing and adverse-selection costs are the most significant spread.

Furthermore, Hasbrouck (1988, 1991a, 1991b) opposes the structural approach, arguing that the inclusion of unobservable variables inevitably results in strong assumptions. Furthermore, he emphasizes the weakness of the model to provide forecasts and underlines that even in the case of a misspecified structural model, a moving average representation, whenever it is possible, could be valid.[67] He proposes MA and AR representations instead, but he recognizes that the choice of the model depends on the nature of the problem. Autoregressive representations offer a powerful analytical tool, especially when it comes to forecasting and impulse functions. Their ability to account for long memories offers a flexible framework to examine the impact of a trade over time and consequently to forecast future values.[68] Such attributes do not exist in the contemporaneous, structural, framework of Roll's (1984) model. However, ARMA representations are not complete, in the sense that they do not fully describe the DGP of returns. If that is the issue, a structural model would be more appropriate. In addition, he recognizes that MA and AR disturbances might be serially dependent, which is in opposition to the structural specification's fundamental assumptions.

---

[65] Traded spreads are estimated first, using the covariance of price change. Then, the adverse-selection component is derived by subtracting the posted from the traded spread. The traded spread is decomposed into the order-processing and the inventory-holding components, which are combinations of estimated probability of trade reversal and the magnitude of a price change as a spread portion (Stoll, 1989).

[66] They find that 77 and 97 percent of the downward bias in previous estimates (Roll, 1984; Stoll, 1989) is caused by time variation in expected returns.

[67] The VAR and VMA (Hasbrouck, 1988, 1991a, 1991b) formulations are based on the Wold theorem that any zero-mean covariance stationary process (x_t) can be represented in the form:

$$x_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j} + k_t,$$

where (x_t) is a zero-mean white noise process, $\theta_0 = 1$, and $\sum_{j=0}^{\infty} \theta_j < \infty$ is a linearly deterministic process. Similar results are discussed in Ansley et al. (1977). In microstructure models, return mean is set to be zero. Although this is a biased estimator of that parameter, its estimation error is lower than the one of the arithmetic mean. For further information see Hasbrouck (2007).

[68] These issues have been further discussed in Hasbrouck (1996) and Dufour and Engle (2000b).

Another influential extension of Roll’s (1984) model was developed by Glosten and Harris (1998). They “re”-introduce the trade initiation variable, which has already been discussed in inventory models and in Glosten (1994), as another source of uncertainty in the formation of the, latent, “fair” value of the asset. They relax the assumption that the random walk component is the only information-related price determinant, and they postulate that trades, along with the liquidity component, carry information that makes investors revise their beliefs. Therefore, they need to take into account in their price setting, the post-trade effect of their transaction. This is consistent with the empirical findings of Hasbrouck (1988, 1991a, 1991b, 1996) and incorporates the dual price impact of trades on a structural framework.

This idea has created a new class of models, which base their inferences on order flow variations; the “trade indicator” models. Madhavan et al. (1997), in an influential study, extend Glosten and Harris’s (1998) approach towards Glosten and Milgrom’s (1985) propositions. They postulate that the informational content of a trade is better measured by the innovation in order flow, considering that the surprise indicates unexpected information. Several other studies extend this model to incorporate the price impact of volume (e.g., Angelidis and Benos, 2009), time (e.g., Grammig et al., 2007; Ben Sita, 2010) or both (e.g., Bowe et al., 2007). In addition, Huang and Stoll (1997) propose a variation of Glosten and Harris’s (1998) “full”, in the sense that it includes all three spread components, model that summarizes many previous approaches, structural and time series, as special cases. Along the same lines, De Jong et al. (1995, 1996), examine closely the relation between the structural model of Glosten (1994) and various VAR specifications, under various transaction sizes.

In contrast, covariance based and trade indicator models have been criticized by other empirical studies, although they confirm the significance of order-processing, inventory-holding and adverse-selection costs as components of the Bid-Ask spread. Clarke and Shastri (2000) find that the Huang and Stoll’s (1997) model provides implausible estimates of the adverse-selection component in 60 percent of cases, for a random sample of 360 NYSE listed stock. Van Ness and Warr (2001), examining how various adverse-selection models are related in measuring information asymmetry, find that, for a sample of 856 NYSE stocks, Huang and Stoll’s (1997) model gives impermissible estimates of the adverse-selection component in more than 50 percent of all cases. Along the same lines, Henker and Martens (2003), using the Huang and Stoll

(1997) model, find a negative average adverse-selection cost component for the stocks of S&P 500 index under both fractional and decimal trading. [69]

Concerning the Carbon market, recently, academic interest has shifted to intraday phenomena, such as the Bid-Ask spread formation. Spread components seem to be of particular interest for regulatory authorities, which are in charge of the design of the upcoming commitment periods. A better analyzed decomposition of the spread components can lead to a more efficient market with smaller spreads. This is of particular interest to operators of exchange platforms, for actively trading investors and agents, such as market makers, as well as for researchers. Only two relevant studies seem to examine the issue. First, Benz and Hengelbrock (2008) have studied the intraday trading process of this young market during Phase I, with respect to price discovery and liquidity. They analyze liquidity and the spread components using the structural model of Madhavan et al. (1997), while price discovery is examined under the VECM (Vector Error Correction Model) framework of Engle and Granger (1987). They find that spreads are always smaller in ECX than in Nordpool, which is also found to contribute to price discovery. Second, Frino et al. (2010) examine the relation between liquidity and trading costs, and find that liquidity has remarkably increased over the past years, resulting in lower spreads.

2.1.5 Further Empirical Issues

Volume

The importance of the empirical models is undeniable, because they address further, empirical, issues in understanding intraday dynamics. One of the most significant issues examined is the price impact of trading volume. Many studies have used transaction volume as an explicit variable to model spread components (e.g., Angelidis and Benos, 2009). Easley and O’Hara (1987) argue that informed traders, under liquidity and optimal strategy considerations, choose either to trade intensively or to strategically segment their trades. Both generate a number of information-motivated trades. In this sense trading volume may convey information, which might have a direct price impact.

[69] These models are build on the idea that in the presence of quote adjustments for inventory management, order flow must be negatively correlated (i.e., Buy (Sell) orders are more likely to be followed by Sell (Buy) orders) as quote adjustments make a transaction on the other side of the market more likely (Serednyakov, 2005). In practice, positive correlation might be observed, and that would be consistent with informed traders splitting their orders into many small orders to disguise their trades. Such behaviour of informed traders does not necessarily mean that liquidity suppliers do not adjust quotes to cover inventory-holding cost when order flow is persistent.

Furthermore, De Jong et al. (1995) focus on transaction costs and show that these costs are lower in small trades, in Paris Burse compared to LSE, and that they are a decreasing function of trading volume. One year later, De Jong et al. (1996), extending Glosten (1994), argue that price changes in Paris Burse increase with trading volume, by 25 percent for small and 60 percent for large trades (De Jong et al., 1996).[70] They use Hasbrouck's (1991a) model and find that the permanent price impact of trading volume varies between 40 percent and 115 percent. Along the same lines, Dufour and Engle (2000b), using again Hasbrouck's (1991a) model, stochastically model time and argue that trading activity is positively associated with volumes. Meanwhile, Pascual et al. (2004) show that trading activity increases after a transaction with information content. Furthermore, Berkman et al. (2005), studying the impact of trading volume on execution costs in the London International Financial Futures and Options Exchange (LIFFOE), establish that effective spreads are smaller in futures than in spot markets, confirming Subrahmanyam (1991a).

Other studies examine the effect of volume on spreads, execution costs and order flows. Chan (2000) points out that, in the Hong Kong Stock Exchange, the information component exceeds the inventory spread component, contrary to Bollen et al. (2004), who, developing a market makers' model, postulates exactly the opposite. In addition, Chan and Fong (2000) indicate that, in NYSE and NASDAQ, trade volume is a more significant variable than the number of trades, though the factors affecting it remain unclear. Moreover, Huang and Stoll (1997) and Ahn et al. (2002) find a reciprocal relation between order-processing cost and trading volume. Degryse (1999) shows that the total trading cost is lower (higher), for small (large) trading volumes, in the Brussels CATS system than in LSE. Moreover, Easley et al. (1997) point out that large trades contain twice as much information as small trades, with Buys and Sells differing only marginally. In contrast, Aitken and Frino (1996) find that Buy orders are associated with larger cost than Sell orders, in the Australian Stock Exchange, due to short-selling restrictions. Their findings confirm those of Chan and Lakonishok (1995). In addition, Hedvall et al. (1997) report that order flow asymmetries depend on the traded volume.[71] More recently, Bowe et al. (2007) examine the price impact of trades in MexDer 28-day interest rate futures contracts, modelling trading volume stochastically, and report that trading activity is inversely related to price changes.

---

[70]

[71] According to their study "*There are more often information reasons behind large Buy trades than large Sell trades*" (Hedvall et al., 1997)

Time

Another branch of the empirical microstructure literature examines the price impact of trading intensity under time considerations. Unlike data of lower frequency, when UHF data sets are employed, time is not set in regular time intervals. Instead, it is irregularly spaced, while time series of “time” appear to be clustered and over-dispersed. Although it took some time for empirical research to widely embrace its importance, the stochastic nature of transaction arrival time has two direct implications. The first is quite straightforward and refers to the fact that several econometric models might not hold, due to heteroskedasticity introduced because of periods of intensive trading activity. Second, following Easley and O’Hara’s (1987, 1992), these fluctuations in trading activity might carry price-relevant information, which, if it is correctly specified, could increase the accuracy of microstructure models.

The importance of duration, defined as the time between consecutive trades, emerges clearly in the studies of Diamond and Verrechia (1987) and Easley and O’Hara (1992). Diamond and Verrechia (1987), using an information-based sequential model, postulate that the motivation of trading in equity markets, lies in information, and especially in the arrival of good or bad news. The implication is that informed traders will always trade upon the arrival of new information. They will buy on good news and sell/short sell on bad news. However, in the case of short-selling restrictions, the price effect of adverse-selection becomes milder, especially with respect to bad news. In contrast, uninformed traders’ trading is considered to be motivated by reasons other than information. Therefore, longer durations are more likely to be associated with bad news.

However, Easley and O’Hara (1992) provide a different explanation for the relation between time and the existence of new information. Following Diamond and Verrechia (1987), uninformed traders’ transactions are considered to be unrelated to the arrival of good or bad news, but informed traders’ actions are seen as being motivated only by new information. An informed trader will choose to transact only if there is new information on the market. Consequently, lower trading activity could be observed due to a decrease in the presence of informed traders, and thus, due to absence of information.[72] In the opposite case, informed traders may tend to trade more frequently upon the arrival of good or bad news, and their presence may quickly be ascertained by

[72] Easley and O’Hara (1992) use the change in participation rate to measure the presence of informed traders.

observing larger trading volumes (Easley and O'Hara, 1987).[73] Therefore, longer durations are associated with absence of price-relevant information. [74]

These two studies differ in their interpretation of trading frequency variations, but they raise two important points. First, they both suggest that time, in the price adjustment process, is not exogenous to information. This is also confirmed by other studies. Hausman Lo and MacKinlay (1992) and Fletcher (1995) introduce duration into their ordered probit models as an explanatory variable, without, though, discussing further the implications of time.[75] Pai and Polasek (1995), although they treat time as an exogenous variable, allow the parameters of their model to depend on duration in simple ways. Similarly, Allen et al. (2005) examine further the information content of durations, when empirically investigating the informational role of the market activity. Second, they emphasize that time homogeneity should not be taken for granted in intraday modelling, since duration might contain price-relevant information. The necessity to consider the time dimension of price process is stressed in the empirical market microstructure literature (Easley and O'Hara, 1992; O'Hara, 1995; Easley et al., 1997) and according to O'Hara (1995) it is an empirical matter. However, no relevant study seems to explicitly model time in the Carbon market.

Drawing on previous studies, Engle and Russell (1998) propose a new simple and convenient model, for modelling inter-trade time intervals, which has gained popularity and has become a benchmark in the empirical microstructure literature. They model duration time-series as dependent point processes, using a new model called Autoregressive Conditional Duration (ACD). The concept is not new and it shares a very strong resemblance with Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models. This triggered a rapid expansion of both theoretical modelling and empirical applications. ACD models consist of a conditional

[73] In that period the term trading volume refers to aggregated transaction size. It is a measure of trading intensity that accounts for volume of trading over a period of time. However, time deformation is not widely modelled yet and any time effects are implicitly incorporated into the analysis.
[74] The above model ascertains that information flow is not continuous and that the absence of informed traders could be interpreted as no new information. However, this can only be valid under the assumption that liquidity is not a hurdle for informed traders to trade at their convenience and that they do not act strategically as in Kyle (1985). O'Hara (2003) points out that liquidity is an important factor for price discovery and hence absence of traders does not necessarily mean absence of information. Traders could also be kept out of the market because of trading halts and/or strategic reasons (Back and Baruck, 2004; Foucault et al., 2003). Market regulators might voluntarily suspend informed trading when the new information is expected to have an extreme impact on price (Diamond and Verrchia, 1987).
[75] The ordered probit model for discrete random variables is used to estimate the conditional probability of trade-to-trade stock price changes on the effects of volume, past price changes and the time between trades.

mean specification and an associated distribution assumption for the DGP of durations. Any formulation that contain past durations is sufficient for the conditional mean, while the associated density function should have a positive support. [76] The initial model examines the linear-ARMA specification and the Exponential and Weibull distributions.

One of the main objections in the literature criticizes the simplicity of the conditional mean specification, in the sense that it does not take into account the asymmetric effects of past durations. Even in the original paper of Engle and Russell (1998) evidence of non-linear effects is reported. The first and quite obvious solution is proposed by Bauwens and Giot (2000), who use a log transformation on both sides of the equation of the conditional mean. This way, short past durations have a decreasing impact on the expected duration, while long durations have an increasing impact. The new model is called Log-ACD. However, Dufour and Engle (2000a) argue that it results in an over-adjustment after very short durations, due to convergence to infinity of the log of zero values. Instead, they propose two other specifications in order to deal with asymmetric effects of past durations in a data-driven way. In the first model, they employ a Box-Cox transformation of the Log-ACD model, namely Box-Cox-ACD (BCACD), allowing the data to determine the size of the effect of past durations. This imposes highly non-linear indirect asymmetric effects on the expected duration. In the second, they extend Nelson's (1991) EGARCH model, and use a piecewise linear parameterization that allows past durations considerably larger or shorter than the mean to have different impacts on the expected duration.

In addition, Jasiak (1998), drawing on the IGARCH model of Bougerol and Picard (1992), and the FIGARCH model of Baillie et al. (1996), conjectures that the linear-ARMA specification fails in capturing duration dynamics in the presence of very long memories. He proposes the Fractionally Integrated ACD (FIACD), which improves significantly the fitting of the initial model. Along the same lines, several other studies account for structural breaks and develop a regime-switching framework. The relevant literature refers to the effect of external events on the expected duration, in a way that different regimes of duration might be present. Zhang et al. (2001) propose the, so-called, Threshold-ACD (TACD), in which different parameters are estimated for different economic events. This is a piecewise specification, where different regimes of

[76] Researchers using the ACD framework, need to specify an equation for the conditional mean along with a distribution. The distribution durations are explicitly assumed to be i.i.d.. Also because of the non-negativity of the durations the potential density functions need to have a positive support. Finally, the chosen distribution determines the Log-Likelihood function that will be maximized.

an relevant variable are allowed to have a different impact on expected durations. Meitz and Teräsvirta (2006) extend the TACD model allowing a smooth transition between regimes. This model is referred to as Smooth Transition ACD (ST-ACD).

Other studies criticize the deterministic character of the conditional mean and propose a "latent-factor"-based alternative. Bauwens and Veredas (2004) propose the Stochastic Conditional Duration (SCD), where the expected duration is modelled as a latent variable. The new innovation series introduce another source of uncertainty. This is a joint distribution model, and its estimation is, computationally demanding.[77] Several studies, such as Liesenfeld and Richard (2003), Ning (2004), Strickland (2006) and Bauwens and Galli (2009), develop this model further, mainly introducing different estimation methods. However, Ghysels et al (2004) argue that the specification of the conditional mean is linked with higher order conditional moments. They ascertain that, not allowing for independent variation of the conditional mean and variance, is a very strong restriction,. Therefore, they propose the Stochastic Volatility Duration (SVD) model, which models jointly duration and volatility. They argue that this is particularly important when intraday liquidity dynamics are of interest. However, this model is not popular due to its complexity.

Moreover, some studies criticize the simplicity of the distribution employed, arguing that it is equally significant to the conditional mean specification and determines the fitting of the model. First, different single density functions have been utilized. The most common distributions used are the Exponential, the Weibull (Engle and Russel, 1998), the Generalized Gamma (Zhang et al., 2001), the Pareto (De Luca and Zuccolotto, 2006) and the Burr (Grammig and Maurer, 2000) distributions. The main findings, concerning the density function employed, indicate that fitting varies considerably, and that depends on the data set employed and on the individual characteristics of the market. A common finding, though, is that increased flexibility is usually rewarded with higher precision in fitting and forecasting.

Furthermore, Bauwens et al. (2004) argue that durations exhibit an idiosyncrasy that cannot be captured by traditional models, especially when a single distribution is

[77] The joint distribution offers an increased flexibility, especially in terms of the hazard functions available. The probability of a transaction to occur given that it has not occurred till now is now influenced by two sources of uncertainty, one for the actual duration and one for the expected. Bauwens and Veredas (2004) use the Kalman filter to estimate the model, which produces consistent and asymptotically normal, but inefficient estimators, since it does not entirely rely on the true likelihood of duration. For further discussion and methods of estimation of SCD please refer to Jacquier et al. (1994).

assumed for the DGP of duration. This gave rise to a new statistical approach, which allows the distribution of durations to be described as a mixture of distributions. This provides researchers with great flexibility, allowing them to account for various issues and stylized facts present in each market. A threshold variable that captures these features can be employed, in order to determine different regimes of duration and/or structural changes. Each one of these regimes can then be associated with a different distribution, or sometimes with different shape of the same distribution.

The models proposed can be divided into two broad categories. In the first, an observable variable can be employed to proxy an economic event that it is assumed to have an impact on durations. De Luca and Zuccoloto (2006) propose a switching regime Pareto ACD model using Terasvirta's logistic smooth transition function, where the shape parameter of a Pareto type-II distribution is allowed to vary across different regimes determined by trading intensity. In the second category, the threshold variable used is a latent, and thus unobservable variable. This way, the unobservable regimes can indicate the presence, or the absence, of factors that cannot be observed but are assumed to have an impact on durations. Some authors have proposed a discrete model (De Luca and Gallo, 2004), while some others have proposed a continuous mixture of distributions (Bauwens and Veredas, 2004; Ghysels et al., 2004).[78] Furthermore, Hujer et al. (2002) propose the Markov-Switching ACD model, which appears to be capable of higher forecasting accuracy. However, these models come at a high computational cost. As a compromise, Hujer and Vuletic (2007) propose the, so-called, Static Mixture ACD model, which shares a strong resemblance to the discrete mixture model proposed by De Luca and Gallo (2004), and provides equally good forecasts. In more detail, they allow the distribution of duration to be a discrete mixture of distributions with regimes determined by a latent variable. Each regime is associated with a different conditional mean equation specification in a way similar to TACD proposed by Zhang et al (2001).

Lately, the idea of copulas has been introduced in duration modelling.[79] Past durations are assumed to come from different distributions that are combined, using a weighting function. Wing (2008) suggests a dynamic copula approach that introduces a time

[78] These authors propose the Stochastic Conditional duration and the Stochastic Volatility Duration models, where expected duration is modelled as a latent factor variable, instead of being deterministic.

[79] Copula is a statistical tool that allows for a flexible formulation of multivariate distributions, accounting for a weighting factor among them. Copulas can combine distributions of different variables or of time series, assuming that previous values come from a different distribution. They are extensively used in options valuation, accounting for various types of risk. For further information please to Nelsen (1999).

varying mixing parameter, which measures the degree of time dependence and describes its structure. Furthermore, De Luca et al. (2008) examine a bivariate and trivariate copula function, comparing it with single distributions of increased flexibility.

Duration and Information

The afore-mentioned studies of Diamond and Verrecchia (1987) and Easley and O'Hara (1992) raise the importance of time in relation to information. Recent ACD studies extend this concept further by examining the connection of time deformation and presence of informed trading. Wong et al. (2009) examine the issue empirically, by utilizing a non-linear, volume enhanced ACD specification. They draw inferences for the Shanghai Stock Exchange, confirming Easley and O'Hara's (1992) propositions about trading activity and information. In addition, Tay et al. (2009) propose an Asymmetric ACD (AACD) to estimate the PIN, assuming a Poisson distribution for the arrival rates of information.

Hujer and Vuletic (2007) develop further this idea by connecting directly duration with type of trading through the associated hazard functions. They conjecture that the instantaneous transaction rates (i.e., the probability that a transaction will occur, given that no transaction has occurred till now) of uninformed traders, since they tend to transact independently of information, tend to be moderately progressive or even constant over time (i.e., not changing, independently of when the last transaction happened). This is a characteristic of the Exponential distribution.[80] In contrast, the presence of informed traders depends on the arrival of news, which is not constant usually and is usually assumed to follow a Poisson distribution. Therefore, the instantaneous transaction rates of their trades should have a different, sharper shape. Consequently, whenever the distribution is Exponential, the regime captures uniformed trading. In contrast, when other distributions, with variant hazard functions, are applied, the associated trades are assumed to be related to information. However, one major drawback of their methodology is that the associated distributions, which determine the shape of the hazard function, are predetermined and not data-driven. Similarly, Gerhard and Hautsch (2007), following Admati and Pfleiderer (1988), conjecture that among uninformed traders there are traders, who observe order flow trying to extract information signals. Then they aggregate relevant information and act according to their individual trading needs (i.e., portfolio composition or risk aversion). This is translated

[80] The hazard function of the Exponential distribution is by definition a flat unity line.

into different trading patterns, which can be described by different hazard functions. Gerhard and Hautsch (2007) connect different shapes with different types of traders.[81]

Moreover, duration has been utilized to proxy several variables when testing market microstructure hypotheses. Easley and O'Hara (1992) postulate that short durations reflect information-based trading. Diamond and Verrechia (1987) postulate that long durations tend to decline prices. Engle (2000) uses duration as a measure of trading intensity and his results confirm that higher trading frequency is associated with information, while longer durations tend to be associated with declining prices. In contrast, Dufour and Engle (2000b) utilize duration as a measure of liquidity. Duration can be interpreted as a natural measure of the speed by which market liquidity is enhanced and prices incorporate new information. Their findings are in line with Easley and O'Hara (1992), in the sense that higher trading activity, measured by shorter durations, is associated with a higher presence of information. Finally, Gourieoux et al. (2004) utilize duration as a measure of risk, in a similar way to Renault and Werker (2002) and Ghysels et al. (2004). In these studies, duration is modelled in a way that captures the risk associated with trading under both price and time uncertainty.

Another similar issue, discussed in the literature, is the time dimension of the price impact of trades and how fast new information is incorporated into prices. Hasbrouck's (1991a) VAR framework provides an excellent analytical tool for examining the after trade effects of a transaction and the time it takes for the information that is revealed by that trade to be fully incorporated into a fully informative price. Dufour and Engle (2000b), analyzing the price impact of trades in a large sample of 18 NYSE frequently traded stocks, use a 5-lags VAR-type model with coefficients varying according to durations. Their results, in line with Hasbrouck (1991a) and Easley and O'Hara (1992), indicate stronger positive autocorrelations of signed trades and larger quote revisions for shorter durations. They postulate that more frequent trading indicates increased information presence and therefore trades have a higher informational price impact. Spierdijk (2004), examining a sample of five NYSE stocks, extends Dufour and Engle's (2000b) findings and shows that large trades increase the trading speed while large returns decrease it. Volatility is also found to be higher for shorter durations. Similar results were reported by Manganelli (2005). Recently, Holder et al. (2004) apply a similar methodology to Dufour and Engle (2000b) in Treasury Note Futures traded in Chicago Board of Trade. They provide empirical evidence that, in contrast to equity

[81] Their model is particularly relevant to this study and it is extensively discussed in Chapter 5.

markets, trade durations are significantly positively related to subsequent returns and the sign of trades.

Duration and Volatility

In parallel, the relation between duration and trading volatility cannot remain unnoticed. The main category duration-enhanced volatility models originated in Engle (2000). He argues that a simple GARCH framework should be insufficient in capturing intraday volatility patterns, if duration is not taken into account. He introduces a class of models, which try to examine issues, such as to predict the next volatility by predicting first the next duration or to test how duration affects the current volatility and whether lagged volatility influences the next duration. More specifically, Engle (2000) combines the ACD and GARCH methodologies into a new model, the ACD-GARCH, or otherwise called UHF-GARCH model.[82] He finds both returns and variances to be negatively influenced by long durations in IBM's stock transactions. His findings are in line with the theoretical predictions of information models, where high trading frequency is expected to have an increased price impact, which would result in higher volatility. Alterations and extensions of the UHF-GARCH model are proposed by Drost and Werker (1996), Ghysels and Jasiak (1998), Grammig and Wellner (2002), Meddahi et al. (2006), Bollerslev et al. (2006), Maller et al. (2008) and Czado and Haug (2009).

The general finding in these models describes a positive relation between return volatility and trading activity. Volatility is found to increase after short durations. This is strong evidence that market participants observe market activity, acquire knowledge and act accordingly. Most studies conclude that when traders observe increased trading activity, they interpret it as information inflow. Consequently, they revise their beliefs and price changes are expected to be more volatile. This is consistent with Easley and O'Hara's (1987, 1992) propositions, which connect trading activity with information.

The price volatility of EUA prices has also been discussed in the Carbon market literature, on an intraday level, as well as on lower frequencies. GARCH-type models seem to be the preferred choice. Earlier studies (inter alia, Benz and Truck, 2006; Paolella and Taschini, 2008) employ simple GARCH specifications on daily data for various energy commodities. Borak et al. (2006) find that price volatility increases

[82] This is an UHF version of the initial GARCH model and along with Realized Volatility models they are the most widely used frameworks in modelling and forecasting intraday volatility. Various studies, such as Kayahan et al. (2002), Corsi et al. (2006, 2008) and Racicot et al. (2008) compare empirically their forecasting accuracy, providing conflicting results.

closer to the maturity date. This contradicts the normal convention that volatility is higher the longer the period of uncertainty and they attribute that to the long life of the EUA futures contracts. Along the same lines, Benz and Truck (2006), employing an AR-GARCH model, postulate that spot volatility affects the long-term investment risk and the efficiency of the market. They also argue that the model outperforms other, constant volatility, models.

Furthermore, other models (see, inter alia, Chevallier, 2009; Vinocur, 2009; Isenegger and Wyss, 2009) investigate the same issue from an intraday perspective. Rittler (2009) employs a GARCH-BEKK model and provides evidence that futures contracts lead the long-run price formation, while in the short-run, bidirectional causality is observed. In parallel, Conrand et al. (2010) employ an Asymmetric Power GARCH and postulate that potential non-linearities, present in the intraday price formation, might not be sufficiently captured by simple GARCH specifications. Further, they underline the importance of NAPs, emphasizing that prior knowledge creates an exploitable information advantage. This contradicts Vinocur (2009) who argues that traders in the Carbon market under-react to new information. Mansanet-Bataller et al. (2010), examining the spread between EUAs and CERs, report that traders in the Carbon market seem not to be influenced by fundamental Carbon related information.[83] Instead, they confirm Viswanthan's (2010) propositions that traders are mainly influenced by microstructure variables. Similar to Conrand et al. (2010), Benz and Truck (2009) model returns and price change variance stochastically, using combination of AR-GARCH and Markov regime-switching models. They conclude that the increased flexibility, in terms of non-linear effects, provides better fitting and forecasting, since it can capture certain characteristics better, such as skewness, excess kurtosis, or structural changes.

Intraday variations

Furthermore, many studies examine the intraday formation of volatility in relation to trading activity, as well as intraday seasonal patterns of other marks. In actual markets, trading occurs in predetermined trading sessions separated by periodical and regular time intervals. Opening and closing times are pre-defined, as well as the lunch break,

[83] Certified Emission Reduction units (CER) and European Union Allowances (EUA) provide the holder with the right to emit one tone of Carbon dioxide. Their main difference is that EUAs are allocated each year by national governments, while CERs can be acquired by an individual company, when it develops a green project. They can be used for the same purpose, which is compliance and their prices should be similar.

the overnight periods, the weekends and the holidays. Several studies investigate the effects of these intervals on intraday variables, as well as their intraday variations. More specifically, Easley and O'Hara (1997) report that trading activity and price volatility, in NYSE and NASDAQ, demonstrate U-shape patterns for liquid stocks. Goodhart and Demos (1991) report similar intraday variations for volume and quoted spreads. In general, well-established markets with well defined opening and closure times, such as NYSE, NASDAQ and LSE, tend to follow similar patterns, while markets, such as the Forex Interbank market, with round-the-clock, partially overlapping trading tend to follow more complex fluctuations (see, inter alia, Wood et al., 1985; Harris, 1986; Dacorogna et al., 1993; Andersen and Bollerslev, 1998).

In addition, weekend returns tend to be lower than the week days' returns (French, 1980; Gibbons and Hess, 1981; Keim and Stambaugh, 1984). Intraday returns and volatility seem to exhibit U-shape patterns (Harris, 1986, 1988, 1989; Gerety and Mulherin, 1994; Andersen and Bollerslev, 1994, 1997; Kleidon and Werner, 1996; Foster and Viswanathan, 1993). Furthermore, open-to-open returns are more volatile than close-to-close returns (see, inter alia, Amihud and Mendelson, 1987, 1988; Stoll and Whaley, 1990; Gerety and Mulherin, 1994). Finally, evidence shows that returns over trading periods are more volatile than returns over non trading periods (Fama, 1965; French and Roll, 1986; Amihud and Mendelson, 1991; Barclay et al., 1990).

Furthermore, Amihud and Mendelson (1987) examine the characteristics of stock returns as reflected by their time series, under the two trading mechanisms of NYSE. They compare the open-to-open and close-to-close returns on the basis of variance ratios of these returns (Amihud and Mendelson, 1991; Stoll and Whaley, 1990). Greater deviations from the random walk form, and therefore from market efficiency, are observed for opening than for closing prices. These differences are attributed to NYSE special characteristics. Other studies find that the variance of the morning call is greater than that of the afternoon call (see, inter alia, Amihud et al., 1990; Stoll and Whaley, 1990; Amihud and Mendelson, 1991), which are attributed to the trading mechanism's special characteristics. Another explanation, though, is given by Leach and Madhavan (1992) and Madhavan (1993) who argue that the higher morning variance is more associated with the overnight closure than with the trading mechanism itself.

Baillie and Bollerslev (1989, 1990) model the dynamic and distributional properties of daily, weekly, fortnightly and monthly returns in the Foreign Exchange market (DM-$).

They estimate GARCH models with t-distributed errors and introduce daily dummies in both the conditional mean and the conditional variance equations. Furthermore, Andersen and Bollerslev (1997) and Andersen et al. (1999) use a flexible Fourier framework into a standard volatility model, in order to model the frequencies corresponding to the different seasonal peaks. [84] Moreover, Dacorogna et al. (1993) use a different time-scale, the J-time, which expands daytimes with high mean volatility and contracts daytimes with low volatility, and seasonal patterns almost vanish with this time scale. Similarly, Muller et al. (1990) and Muller and Sgier (1992) investigate time patterns in the Forex market and account for the time dimension of the global market activity as a combination of regional time patterns.

The intraday patterns of other components have also been examined. Several studies focus on the intraday seasonality of the Bid-Ask spread components. Brockman and Chung (1998) and Chan (2000), studying the intraday price formation of the Hong Kong Stock Exchange, report that the spread components and their impact on price formation tend to follow a U-shape pattern. Menyah and Paudyal (2000) examine the spread components in the London Stock Exchange. They find that the inventory cost component is smaller than the order-processing, while the asymmetry information component accounts for approximately 47 percent of the quoted spreads. Kim and Ogden (1996) report a similar proportion for the adverse-selection component in NYSE/AMEX stocks. In contrast, Declerck (2000), examining the trading costs and their components in CAC 40 (i.e., Cotation Assistée en Continu) index stocks, reports that the order-processing cost explains 82 percent of the spread, while all components are found to be positively related to trading volume. Similar to Kim and Ogden (1996) and Menyah and Paudya (2000), Silva and Chavez (2002) report that the higher execution costs of the Mexican Stock Exchange can be attributed to a higher adverse-selection cost. Furthermore, Ahn et al. (2002) provide evidence of order handling cost components that follow U-shape intraday patterns in the Tokyo Stock exchange.

In a similar study, of the SIMEX futures contracts on the NIKKEI 225 index, Kim et al. (2002) report an L-shape pattern for asymmetry information cost component and an inverse U-shape pattern for inventory-holding costs. Gwilym and Thomas (2002)

[84] *"A Fourier series is an expansion of a periodic function f(x) in terms of an infinite sum of sines and cosines. Fourier series make use of the orthogonality relationships of the sine and cosine functions. The computation of these series (harmonic analysis) is useful in breaking up an arbitrary periodic function"* , in this case intraday patterns, *"into a set of simple terms that can be plugged in, solved individually and then recombined to obtain the solution to the original problem or an approximation to it to whatever accuracy is desired or practical"*. (http://mathworld.wolfram.com/FourierSeries.html

examine several spread measures using trades and quotes of the LIFFE FTSE 100 index of futures contracts and reports wider spreads on the opening and narrower on the closing of the trading day. Transaction spreads are found to be biased estimators of the quoted ones. Furthermore, intraday seasonality of spreads has been studied in options markets as well. Norden (2003), using data from the Swedish OM market, provides evidence of both in- and out-of-the-money options being more asymmetric than at-the-money ones. He also reports that the theoretical option value is closer to the Bid. Pinder (2003), examining spread components in the Australian Options Market, before and after a switch from a quote-driven floor-traded to an order-driven screen-traded system, reports narrower spreads when quoted prices are continuously provided.

More recently, Ben Sita (2010), using data from Helsinki Stock Exchange and NYSE, argue that the Bid-Ask spread in a pure limit order book market contains a risk component associated with managing time, which accounts for 19.6 percent, and that the adverse-selection cost exhibits a U-shape pattern. Angelidis and Benos (2009), report that adverse-selection cost exhibits a U-shape pattern in Athens Stock Exchange (ASE). The cost component follows a U-shape curve for high priced stocks and it is a function of time for lower priced stocks. Consequently, the vast majority of the studies concludes that the intraday behaviour of all variables of interest exhibits strong seasonal patterns that need to be modelled accordingly.

Concerning the Carbon market, only two relevant studies seem to exist. First, Benz and Hengelbrock (2008) examine intraday spread formation, using Madhavan et al. (2007) model. They report that the only trade related price determinant is information, while the liquidity component is small and rather insignificant. The information component is found to significantly vary in the more liquid ECX, which is also found to be the price leader in intraday price formation. Along the same lines, Frino et al. (2010) examine the relation between liquidity and trading costs, and report that spreads decrease as market gains complexity and liquidity increases. In addition, they examine the impact of informed trading on prices, spreads and volatility. Their findings confirm previous literature, providing evidence of an inverse relation. Higher presence of informed traders results in wider spreads.

Learning, Type of order and Liquidity

Two issues referring to the behaviour of market participants are particularly relevant to the discussion of the following chapters. First, several studies recognize that market participants observe and learn from the actions of others. In the inventory models (inter alia, Amihud and Mendelson, 1980; Stoll, 1978) dealers try to extract information by observing past trading activity and then formulate expectations concerning potential incoming order flow imbalances. In the information models (see, inter alia, Kyle, 1985; Glosten and Milgrom, 1985; Easley and O'Hara, 1987, 1992) they aggregate information and revise their beliefs for the "fair" value of the asset. However, several phenomena occur because different traders interpret differently the same information signal and they probably have different learning curves.

The actual definition of learning and the speed of information dissemination have been extensively discussed in previous studies.[85] Market participants observe the same information (i.e., trading process), but they might interpret order flow signals in a different way, because of factors related to their prior knowledge or their market share. In addition, their access to information and the speed of disseminating it varies considerably and therefore a more dynamic approach should be more relevant. REE literature (see, inter alia, Shiller, 1981, 1984; De Long et al., 1990; Chamley, 2003; Sandroni, 2005) recognizes that learning is a dynamic process and that apart from the "actual" truth, there are various signals that might turn traders towards the "right" or the "wrong" direction (see, inter alia, Banerjee, 1992; Bikhchandani et al., 1992; Smith and Sorensen, 2000). "Actions speak louder than words" and therefore traders might "herd" towards an unexpected signal. However, according to Vives (1993) and Jun and Vives (2004), the way they interpret this signal, as well as the speed in which they accumulate and disseminate information, might vary considerably.

The second issue refers to the price impact of different types of orders. As Limit Order Markets (LOMs) have developed, many studies discuss the informational efficiency of these markets (e.g., Brown and Zhang, 1997), the relation between type of orders and traders or spreads (e.g., O'Hara and Oldfield, 1986; Chakravaty and Holden, 1995) or more importantly the trader's choice concerning the type of trades (see, inter alia, Cohen et al., 1981; Angel, 1994; Kumar and Seppi, 1994; Harris and Hasbrouck, 1996; Harris, 1998; Parlour, 1998; Foucault, 1999; Goettler et al., 2005; Wald and Horrigan, 2005). More recently, Bouchaud et al. (2004, 2006) and Wyart et al. (2008) derive the lagged impact function from the structural model of Madhavan et al. (1997) and conjecture that the main determinant of the spread is the adverse-selection component

[85] This group of studies belongs to the REE literature, which is beyond the scope of this thesis. Therefore, only some indicative studies will be discussed here.

while liquidity causes most of the volatility. Their derivation helps to determine the marginal profit of limit and market orders. However, all studies seem to agree that limit orders are associated with wider spreads.

The majority of studies deals with high liquid stocks in well-established markets. Some recent studies, however, examine less liquid, emerging and new markets. Landstiel (2005) studies the market microstructure of illiquid option markets and their interrelations with the underlying spot markets. Decomposing the spread to order-processing, δ hedging, inventory-holding costs and competition, he shows that the spread depends positively on δ hedging cost and the spread of the underlying security. More recently, Angelidis and Benos (2009), extending on the well-known Madhavan (1997) model, analyze the components of the spread in the Athens Stock Exchange (ASE), by introducing trading volume. They estimate the spread components for large and medium capitalization stocks and find that the adverse-selection component follows a U-shape pattern and that all cost components depend on the stock price. Bowe et al. (2007) focus on the price impact of duration and trading volume in the emerging futures market of Mexico (MexDer 28-day interest rate futures contracts). They find that the duration between transactions exerts a positive influence on price changes, similar to order flow innovations. In contrast, trade volume affects prices negatively. They conclude that the dominance of the liquidity component suggests that liquidity traders dictate the time to trade.

2.2 Reflections on Literature

Price formation, unlike the macroeconomic approach, on an intraday level, is considered to be driven by information and liquidity. Information changes investors' beliefs about the "fair" value of the asset and therefore it moves prices permanently. In contrast, liquidity variations might affect market participants' exposure, especially on a dealer assisted environment such as the Carbon market. Large order flow imbalances might force them to deviate from their optimal portfolio positions. As a result, they will adjust their price setting in order to build or decrease their inventory. Therefore, that effect is only transitory and it lasts only until they return to their desired position. Consequently, information and liquidity are seen as the main determinants of the Bid-Ask spread as well. Market makers will on average lose money when they transact with better informed traders. However, they can compensate by charging a higher spread and achieving higher profits from trading with uninformed traders. Along the same lines, liquidity variations have a direct impact on trading costs, especially the cost of providing "immediacy", where spreads might widen when trading activity is low.

The literature recognizes two types of trade, and consequently associated traders, with respect to information. First, informed traders possess price-relevant information prior to other traders and they have an incentive to exploit it. Uninformed investors, however, trade for liquidity reasons, unrelated to information. These traders can be further divided according to whether they can choose the optimal time or size of transaction, in order to maximize their profits. Some uninformed traders are considered to observe the market trying to extract information. This way, they try to partially exploit the informational advantage of informed traders before it becomes public knowledge. This implicitly assumes three things. First, there is a period of time in which information is exploitable, from the moment it hits the market up until it becomes public knowledge. Second, a different learning process is assumed for different market participants, both in terms of quality and speed of learning. Third, investors observe specific market variables, trying to extract price-relevant information from their variations.

Transaction size and trading frequency have been extensively discussed in the literature. Inter-trade duration is irregularly spaced and higher transaction arrival rate is seen as a sign of information presence. Similarly, transaction size has long been utilized as a key concept in identifying the informational content of transactions. Increased transaction size and/or shorter durations are seen as indicators of increased presence of informed trading. Therefore, investors can observe variations in intensity of trading, measured by both size and time, and they can draw inferences on whether there is still price-unresolved information. In addition, these trades are recognized to have a more significant price impact. In practice, this means wider price variations in active stages of the market and consequently increased volatility and risk. Furthermore, both trading volume and frequency are recognized to be determinants of the Bid-Ask spread.

Concerning the Carbon market, researchers have only recently shown an interest and in the last couple of years an increasing number of studies investigate various aspects of regulation, structure and trading activity. Early literature, especially in Phase I, is quite sparse and deals mainly with describing the legal framework, the origins and the structure of the market. However, academic interest shifts to more empirical issues, such as the price impact of banking restrictions and allocation of allowances. In addition, the economic nature of the asset, as well as the relation between spot and future prices, or other commodities, has been further examined. More recently, the investigation of price volatility has gained some attention. Several studies use the UHF-GARCH framework to investigate intraday price and volatility formation, as well as several other stylized facts. Furthermore, UHF-data provide the opportunity for examining further the microstructure issues, such as price leadership between exchanges or alternative products (EUAs and CERs), or simply intraday price discovery. Some microstructure issues can be further examined. For example, to my knowledge there is no relevant study, modelling duration in the EU ETS or in the Carbon market in general. In addition, spreads have not been investigated extensively, with only two relevant published studies, while intraday price formation is dominated by GARCH-type time series models. Structural approaches can be investigated further, since there is only one relevant (Benz and Hengelbrock, 2008).

Drawing on the relevant literature, the present study emphasizes several issues that they can be developed further. First, the scarcity of empirical studies in the Carbon market provides an excellent opportunity for further issues to be discussed. The unique characteristics of the market, along with the absence of relevant studies, make the investigation of time deformation particularly important. In a politically influenced, “cap and trade” environment, in which market innovation is supported by allowing OTC EUA holders to enter the market, duration modelling might reveal further insights of the trading process. Second, the simultaneous modelling of transaction size and time provides the opportunity to examine their combined effect. Emphasizing the informational content of trades, previous studies consider them as indicators of subsequent price changes. In addition, the hazard function of duration has been previously used to identify different types trades with respect to information. Therefore, the implementation of duration analysis with trading intensity could provide a more accurate depiction of informed trading. Third, several studies have examined the learning process, in terms of quality and speed, but their approach is mostly theoretical. A empirical approach could reveal further structural aspects. Fourth, it would is important to examine how trading history affects traders’ expectations and consequently prices, when trades can carry information, which can be extracted by discretionary uninformed traders. Structural models recognize that prices incorporate investors’ expectations and, therefore, provide a solid foundation for that analysis. Finally, this idea can be further developed to account for the dual character of the information that can be extracted from order flow. Investors can formulate expectations regarding the "fair" value of the asset, as well as the future levels of liquidity. This could be achieved by utilizing a dynamic approach in formulating expectations, jointly modelled with price.

Chapter 3

Data

3. Data

Chapter 3 presents the data employed for the analysis in the following three chapters, as well as a preliminary analysis of various market variables. Section 3.1 presents the data collection process as well as the preparation of the specific data sets for the following intraday analysis. Section 3.2 presents a preliminary, non-parametric analysis of price, duration and trading volume formation over the years and on an intraday basis.

3.1 Data Collection and Preparation

The data employed in this study concerns the two largest exchanges, namely ECX and Nord Pool, of the EU ETS market. The data sets cover the period from market inception, namely January 2005, till the end of 2008.[86] This period includes the whole Phase I and the first year of Phase II. These phases are examined separately in each market. The data consists of all recorded transactions of futures contracts. The information, marks, reported for every transaction, includes date, time stamp, price, volume, direction of trade and an indication of whether a transaction was OTC.

The particular asset of interest is the European Union Allowances (EUAs) futures contracts with maturity date December 2008. These contracts have been chosen because they are by far the most liquid contracts and for consistency with previous literature.[87] In more detail, all previous, relevant studies employ future contracts, instead of spot prices, due to liquidity reasons and in order to account for the impact of banking restrictions in between phase I and II. In addition, these studies focus on futures contracts with maturity date December 2008, because the liquidity levels of the other contracts (futures with different maturity and options) is restrictive for an intraday analysis.

Futures are standardized contracts that give the right, with the symmetric obligation, to their holder to buy or sell a certain amount of EUAs at a predetermined future date at a predetermined price. Every contract, lot, corresponds to 1000 EUAs and every EUA gives the right to emit 1 ton of CO_2 equivalent greenhouse gases. The settlement of the

[86] 2005 was the first year of operations for the EU ETS and as the market was in a very early stage and rather unstable, all observations of that year are omitted.

[87] The precise maturity date is the first business day of December on Nord Pool and the last Monday of December on ECX. These contracts can be used for compliance reasons on April 2009. Regulated companies that they need to comply for their emission for 2008, can use futures contracts either for hedging or for speculation. For further information refer to www.ecx.eu

contracts is guaranteed by the respective clearing house, while counterparty risk is mitigated by a margin account.[88] Prices in both exchanges are quoted in Euros and the minimum tick is €0.01. Trading is continuous from Monday to Friday, with trading hours 08:00-18:00 Central European Time (CET) on ECX and 08:00-15:30 on Nord Pool).

In addition, the microstructure literature poses some issues concerning data manipulation that need to be taken into account. First, all transactions out of the official trading hours are excluded, since only trading patterns within the normal continuous trading period are to be examined. These are observations before 08:00 (08:00) or after 18:00 (15:30) in ECX (Nord Pool). Second, the data is organized as a continuous trading session. This is done in order to account for intraday patterns over a period of time. Durations, similar to Ben Sita (2010), are calculated in seconds and the overnight exchange close period is treated as if it does not exist, in order to avoid heteroskedasticity of known form. This, contrary to Manganelli (2005), implicitly assumes that there is no price-relevant information revelation during this time.[89] Third, in order to deal with the asymptotic convergence to minus infinity at zero of the logarithmic function, all zero durations are omitted and all associated "marks" are aggregated into the first transaction reported.[90]

Another important issue is the treatment of outliers. Phase I was the pilot period for the EU ETS and some unusual observations, such as extremely long durations or high volumes, are observed. In addition, the construction of continuous trading data sets that ignore non-trading periods creates some artificial observations, such as durations longer than the official trading hours. Therefore, the following filters are applied. First, all

---

[88] The ICE Clear Europe, clearing fee is €3.50 and €3.00 per lot per side in ECX and Nord Pool, respectively.

[89] For example, the time elapsing between 16:59:30 of day t-1 and 07:00:10 of day t is considered to be only 40 seconds. The same rule is applied in all days without transactions, such as weekends and holidays. They are treated as if they do not exist. There is a great debate on the implications of either including or excluding these time intervals. More specifically, authors like Ben Sita (2010) maintain that, when non-trading periods, such as weekends, are included in the data sets, heteroskedasticity of known form is imported because of the seasonality involved. On the contrary, Manganelli (2005) argues that the elimination of the overnight period results in the loss of important information.

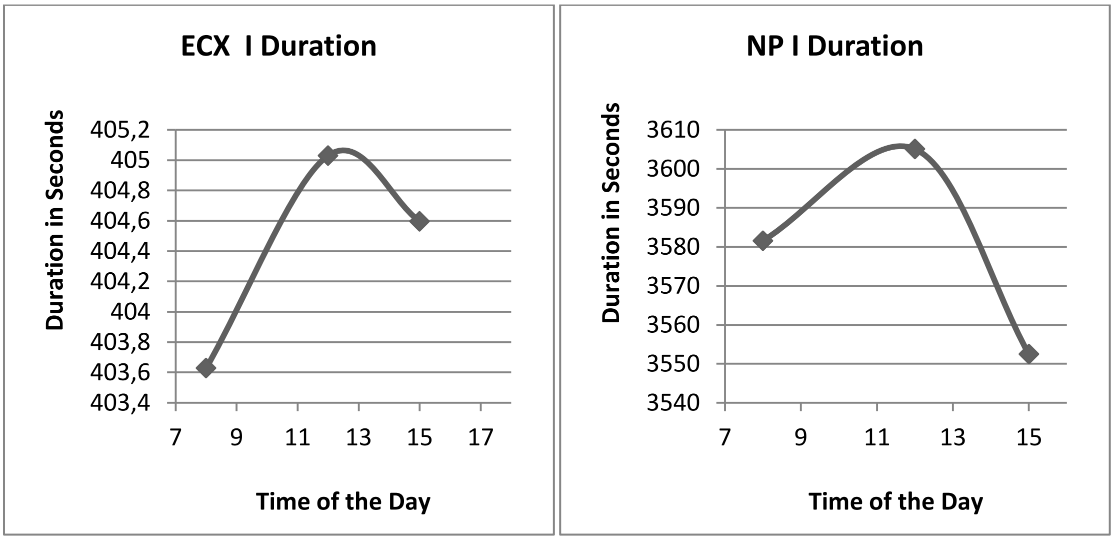
[90] The term "aggregation" refers to volume, where the value used in the final data set is the sum of all relevant values from the omitted transactions. For example, assume that there are four transactions with the same time stamp, where the associated volume for each one is five contracts. In the final data set, these four transactions would be included as one, single transaction. The volume of this transaction would be 20 contracts. In addition, the price, the trade sign and the dummy variable that picks OTC transactions are also affected. However, the majority (over 90 percent) of these transactions have the similar figures. For example, when there are many transactions with the same time stamp, in 90 percent of the cases, these transactions have the same price, the same trade direction and they are of the same type (i.e., OTC or non-OTC). Therefore, only the relevant marks of the first transaction are taken into account.

observations, with durations longer than the official trading period, are omitted. Second, all observations, with durations longer than the mean plus five standard deviations that can be considered outliers are omitted. The same procedure is applied to price as well. Finally, all observations, with volumes larger than 500 contracts, are omitted to account for recording discreteness. This filtering procedure generates four data sets.

Phase I In sample 1/2/2006-31/10/2007	ECX 42,606 observations
	Nord Pool 3,804 observations
Phase II In sample 1/2/2008-31/10/2008	ECX 91,264 observations
	Nord Pool 3,606 observations

Finally, the vast majority of the microstructure literature reports a strong intraday trading seasonality, with markets being more active than average immediately after the opening and just before closing. This inserts heteroskedasticity into duration and trading intensity time series and therefore needs to be dealt with. Figure 3.4, below presents the intraday variations of inter-trade durations in both markets and phases. All four panels indicate that duration exhibits the usual inverse U-shape intraday pattern in both markets and phases. Market activity is more intense during the opening and closing sessions, while duration is notably longer during the lunch break.

Figure 3.1: Intraday seasonal pattern of Duration



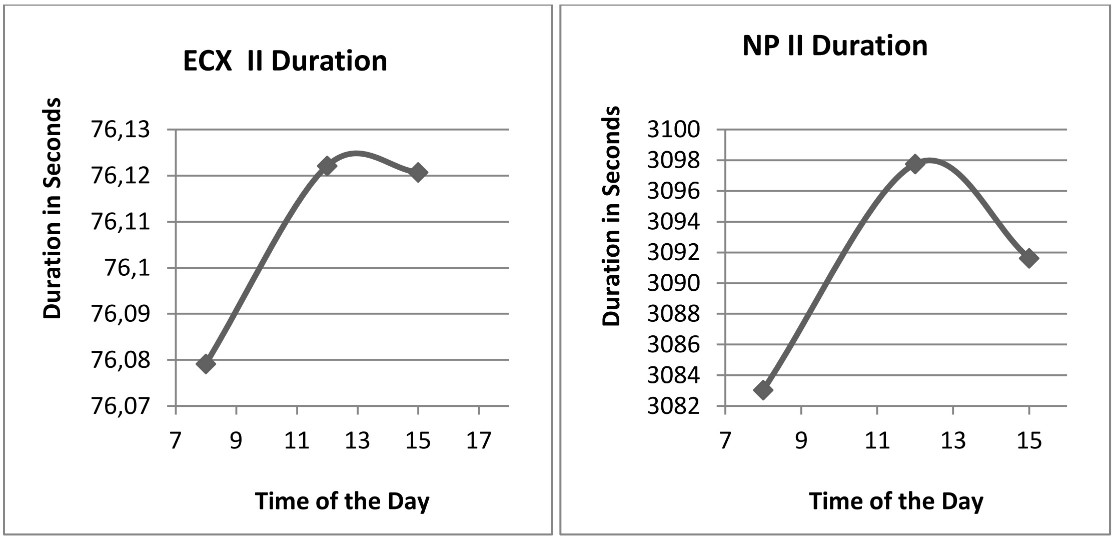


Figure 3.4 presents the intraday patterns of actual durations and trading intensity in both markets and phases.

Engle (2000) uses a diurnal adjustment of durations, which is applied to the duration series in this study. Trading intensity is also de-seasonalized. Briefly, raw values of duration and trading intensity are regressed on a cubic spline function of the daily trading time. The new values are then divided by the expected, modelled on time, values. Time series of this ratio is taken as the diurnally adjusted series.

Specifically, each trading day is divided into five time intervals, each two hours long. The nodes, or time benchmarks, used, are 10:00:00, 12:00:00, 14:00:00, 16:00:00, 18:00:00, which represent 36,000, 43,200, 50,400, 57,600 and 64,800 seconds after midnight, respectively. Then, raw durations ($d_i = t_i - t_{i-1}$) and raw trading intensity, $k_i = {}^{v_i}/_{d_i}$, where v_i is the number of contracts per transactions, are regressed on the following time function, in order to obtain $E(d_i|f(t))$ and $E(k_i|f(t))$.

$$f(t) = \beta_0 + \sum_{m=1}^{3}\sum_{j=1}^{5}\left(\beta_j\left(t - n_j\right)^m\right), \qquad (2.1)$$

where $j = 1, \dots, 5$ stands for the five nodes used, $m = 1, \dots, 3$ and n_j's are five dummy variables, calculated as:

$$n_j = \begin{cases} t - k_j & when\ k_{j-1} < t < k_j \\ 0 & elswhere \end{cases}. \qquad (2.2)$$

Then, after estimating β_0 and the β_j's, durations and trading intensity are normalized, and thus diurnally adjusted, as follows.

$$x_i = {d_i}\Big/{E(d_i|f(t))}, \quad (2.3.1)$$

$$S_i = {k_i}\Big/{E(k_i|f(t))}, \quad (2.3.2)$$

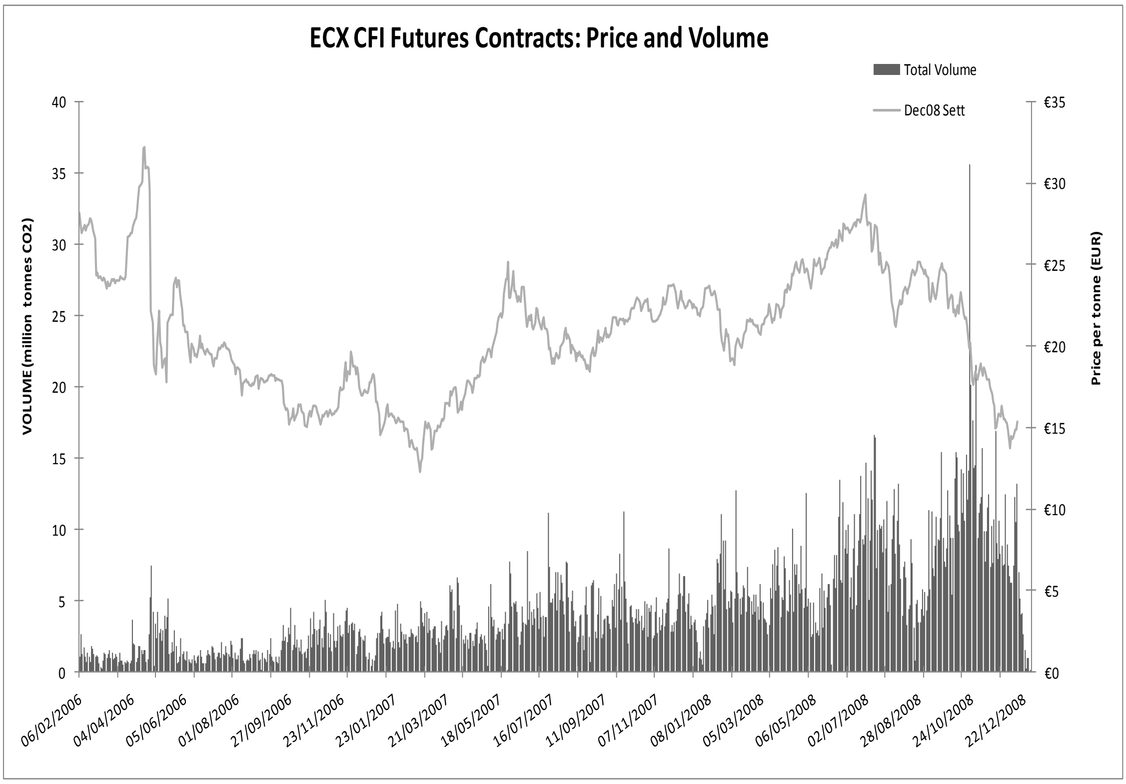
where x_i is the diurnally adjusted durations and S_i is the diurnally adjusted trading intensity.

3.2 Preliminary, Non-Parametric Analysis

Before the main empirical analysis, which is presented in the following three chapters, this section presents some preliminary features of the data series under investigation. The trends over time of the main variables are discussed first. This is followed by an examination of the unconditional. basic statistics. The analytical focus is upon potential differences between the two markets, namely ECX and Nord Pool, and between phase I and phase II.

Figure 3.5 presents daily prices and daily aggregated volumes from the beginning of the market in 2005 till the end of the first year of Phase II, in both markets, namely ECX and NP. These two graphs reveal some interesting features of the market. First, an increasing trend in market activity is observed throughout the time period plotted. The total trading volume increases constantly and seems to exhibit seasonal patterns. Trading activity seems to reach high peaks during the summer and at the year end, while low peaks are observed at the beginning of calendar years. A partial explanation could be related to the weather and the operating cycle of energy-intensive, regulated companies. In addition, the total number of traded contracts decreases as maturity approaches. This is to be expected, since, in case of compliance, the information advantage of futures contracts expires, since traders can buy spot contracts directly with considerably lower exposure risk. Probably for similar reasons, price decreases as maturity approaches. In addition, the average price remains under €25. Considering that the aim of the market is to make energy consumption expensive, a relatively low price per tonne of CO_2 raises concerns about the over-allocation of allowances that results in a relatively low price, especially in Phase I (e.g., Daskalakis et al., 2006, 2009; Mansanet-Bataller and Pardo, 2008)

Figure 3.2: Trade Price and Volume



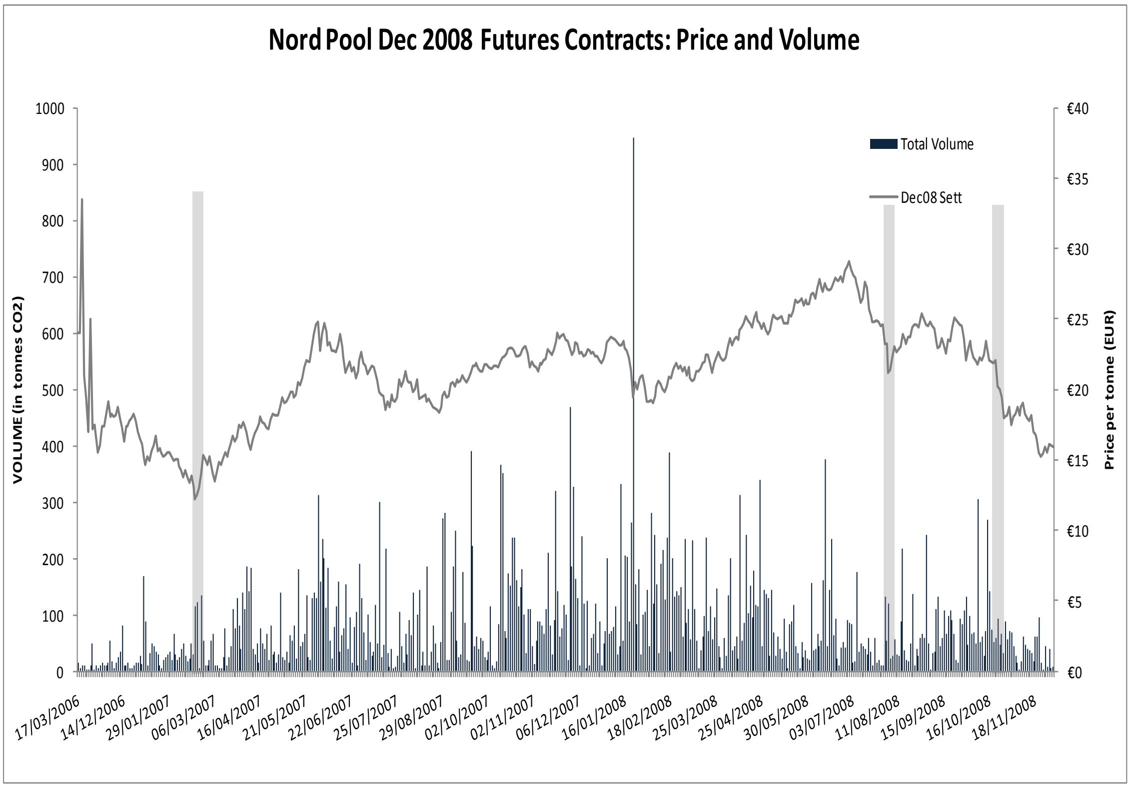


Figure 3.5 presents the daily prices and aggregated volumes across markets and phases

Another observation refers to the two sharp price drops. The first is observed in the beginning of 2006, while the second is observed at the end of 2008. In March 2006, Spain announced an excessive surplus of allowances, due to conservative expectations of emissions. Following the announcement of its NAP, markets interpreted this as a sign of increasing supply. This caused futures contract prices to fall drastically. As it appears in the graph, this unexpected announcement introduced uncertainty into the market, which had a long lasting effect. Carbon prices remained under 20 Euro for long, while signs of recovery appeared a year later. Similarly, in October 2008, two months before the maturity date, after the collapse of Lehman Brothers, economic growth slowed down universally. The Carbon market followed all other markets and the credit crisis decreased EUA prices sharply.

Emphasizing the trading process in the market, Figure 3.6 presents the progression of duration, volume and trading intensity from the inception of the market until the end of the first year of Phase II. The first panel of Figure 3.6 presents the average duration in both markets and phases. The second panel reports the average diurnally adjusted duration and trading intensity, as well as the average volume, in ECX in both phases. The third panel presents the same statistics for NP.

Figure 3.6 confirms an increasing trading activity over the years. The average duration, defined as the time between consecutive trades, decreases drastically over the years, indicating an increased number of transactions per unit of time. This is consistent with the increased trading volume and is confirmed by the next two panels. These show that trading intensity fluctuates a lot, but it follows an increasing trend towards the end of the life of the contracts. A closer inspection of these graphs reveals that the increased intraday trading activity is mainly caused by higher trading frequency, in the form of shorter durations, and not by higher trading volume. In fact, the average transaction size decreases.

Table 3.1 reports the basic statistics of the variables under examination. The first panel of Table 3.1 reports the mean, the median, the max and min, the standard deviation, the skewness and the kurtosis of the actual duration (in seconds), trading volume (contracts per transaction), trading intensity (volume over duration) and price (in Euros), as well as of the diurnally adjusted duration and trading intensity. This panel is divided into four sections one for each market and phase, namely, ECX I, ECX II, NP I and NP II. The second panel of Table 3.1 presents the average and the standard deviation of the same variables, in both markets per quarter. The first section of that panel refers to ECX, while the second refers to NP.

Figure 3.3: Duration, Volume and Trading Intensity

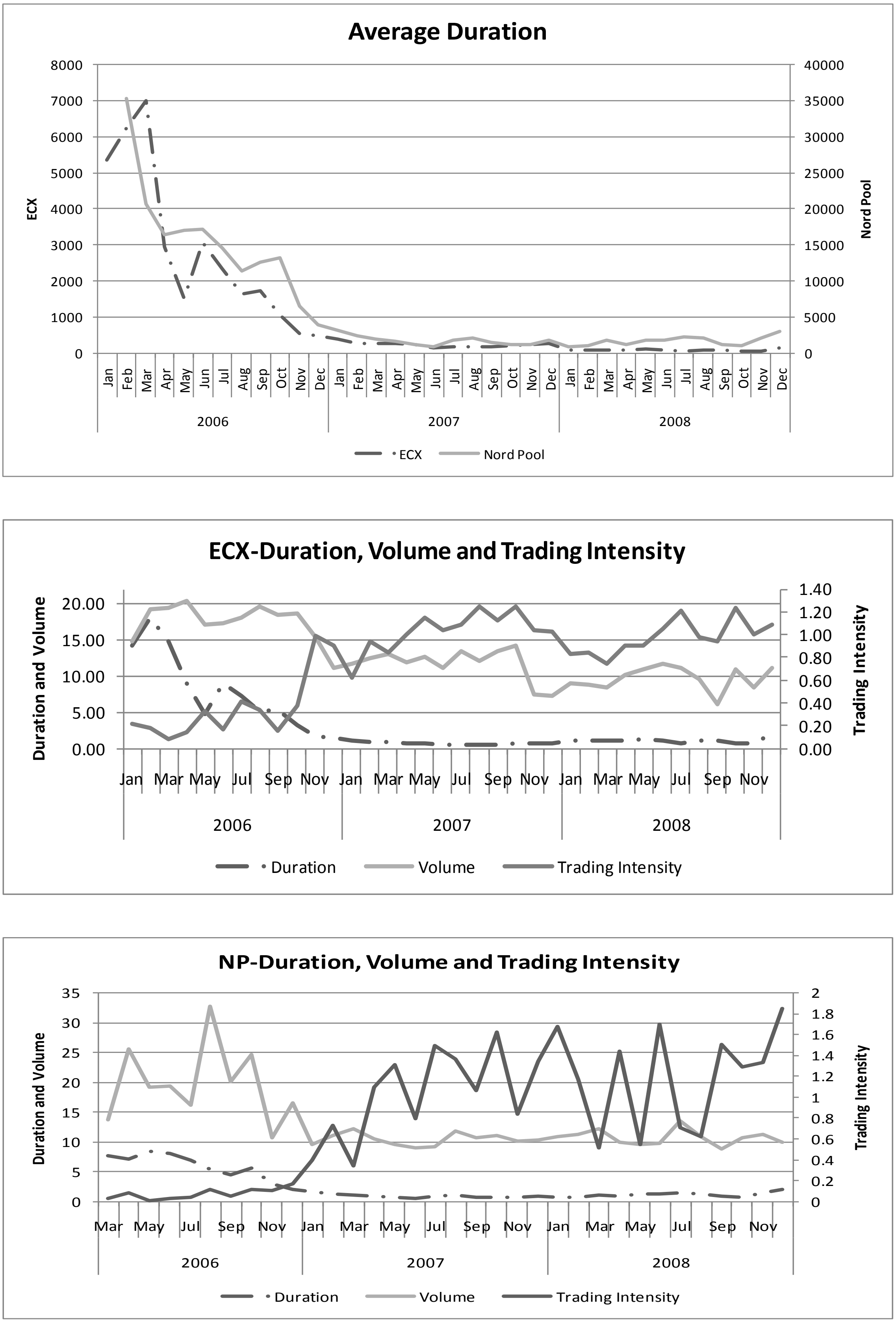


Figure 3.6 presents Duration (diurnally adjusted), Volume (transaction size) and Trading Intensity (diurnally adjusted) across markets and phases.

A quick inspection of Table 3.1 reveals that the two markets differ significantly. ECX is far more liquid than NP with the average duration in both phases being considerably shorter (e.g., 375.87 in ECX I, 75.73 in ECX II, 1,912.21 in NP I and 1,370.57 in NP II). These statistics indicate that liquidity is an important issue between Phase I and Phase II. As market gains complexity and maturity it becomes more liquid as well. The average duration in Phase II is considerably shorter than in Phase I. But the average volume, measured as the average number of contracts per transaction, appears to be fairly similar in both markets and phases (e.g., 12.45 in ECX I, 9.69 in ECX II, 10.81 in NP I and 10.67 in NP II). However, the aggregated figures reported in Figure 3.5 show that ECX is more active and that volume increases over time. This indicates that the increased trading activity over the years and especially in ECX, observed in the second panel of Table 3.1 (diurnally adjusted duration progresses from 0.10 to 1.28 in ECX and from 0.01 to 0.37 (in Q2) in NP), is mainly caused by higher trading frequency. Along the same lines, prices are fairly similar in both markets in every phase. This pattern is expected since the assets traded are almost identical, apart from the date to maturity.

The standard deviation of price is considerably lower than the mean (e.g., the mean and the standard deviation of price are 19.85 and 2.97 in ECX I, 22.60 and 3.33 in ECX II, 20.33 and 3.00 in NP I and 22.73 and 2.80 in NP II). This, along with the considerably low values for skewness and kurtosis (the skewness and kurtosis of price are 0.37 and 2.74 in ECX I, 0.41 and 2.65 in ECX II, 0.64 and 2.78 in NP I and 0.35 and 4.47 in NP II), indicates a distribution with observations closely scattered around the mean.

However, duration, volume and trading intensity appear to be over-dispersed. The standard deviation of duration is two to four times larger than the mean (e.g., the mean and the standard deviation of duration is 375.87 and 1,264.1 in ECX I, 75.73 and 157.15 in ECX II, 1,912.21 and 3,727.12 in NP I and 1,370.57 and 2,195.64 in NP II). This indicates a distribution with positive support and long right tails. In addition, the relatively small mean, and the large maximum values (e.g., the mean and the maximum values of duration are 375.87 and 30,979 in ECX I, 75.73 and 4,332 in ECX II, 1,912.21 and 29,933 in NP I and 1,370.57 and 20,542 in NP II) indicate that observations are gathered around the mean, which is closer to zero, and that the right tail is indeed long. This effect is milder in Phase II, due to increased trading activity, but it is still noticeable.

A closer inspection of the the values for skewness (e.g., the skewness of duration is 11 in ECX I, 7.29 in ECX II, 3.71 in NP I and 3.28 in NP II), indicates that the distributions of the variables in ECX have a longer right tail than in NP. Trading activity is higher in ECX and therefore durations are more concentrated closer to zero, than in the less liquid NP, and thus the higher values for skewness are observed. Furthermore, the median of duration and volume (e.g., the median of duration and volume is 79 and 10 in ECX I, 23 and 5 in ECX II, 480 and 10 in NP I and 540 and 10 in NP II) reveals that ECX is far more liquid than NP and that Phase II is more active than Phase I. In Phase II the average duration is shorter in both markets and according to the standard deviation observations are closer to the mean. Also, the relative values for the skewness and kurtosis are smaller, indicating less frequent and large values.

These findings are confirmed in Figure 3.7, which presents the histograms of duration and trading intensity in both markets and phases. The first four panels of Figure 3.7 present the distributions of durations, while the next four panels show the distributions of trading intensity in ECX I, ECX II, NP I and NP II. Similar to previous findings, the observations of duration are gathered closer to the mean in ECX than in NP, which results in longer right tails. In addition, the observations of trading intensity, in the last four panels in the bottom of Figure 3.7, are closer to the mean, and thus closer to zero, in Phase II than in Phase I.

The last section of Table 3.1 shows the time trends of the variables under examination. Duration decreases (e.g., from 5,191.40 to 66.05 in ECX and from 15,413.75 to 1327.78 in NP) over the years. Over-dispersion also decreases. Volume also decreases over the years, but, according to the standard deviation, is more scattered, as the market gains liquidity and complexity (e.g., the mean and the standard deviation of volume change from 17.88 and 18.38 to 9.86 and 23.29 in ECX and from 13.75 and 11.09 to 10.82 and 17.66 in NP). Trading intensity, however, increases and becomes more concentrated (e.g., the mean and the standard deviation of trading intensity change from 0.10 and 0.65 to 1.28 and 5.09 in ECX and from 0.01 and 0.01 to 0.18 and 0.93 in NP). This confirms that trading activity mainly increases due to higher trading frequencies rather than higher transaction size. Last but not least, prices appear to follow a decreasing trend and volatility appears to be higher closer to the maturity date (e.g., the mean and the standard deviation of price change from 24.58 and 1.95 to 18.83 and 2.71 in ECX and from 24.08 and 0.10 to 19.97 and 2.56 in NP). This could be attributed either to the credit crisis, or to the unique features of the underlying asset.

Table 3.1: Basic Statistics

A.

	ECX I						ECX II					
	Actual Duration (Seconds)	Actual Volume (No of Contracts)	Actual Trading Intensity	Price (In Euros)	Diurnally Adjusted Duration	Diurnally Adjusted Trading Intensity	Actual Duration (Seconds)	Actual Volume (No of Contracts)	Actual Trading Intensity	Price (In Euros)	Diurnally Adjusted Duration	Diurnally Adjusted Trading Intensity
Mean	375.87	12.45	0.69	19.85	1.16	0.98	75.73	9.69	1.12	22.60	1.01	1.01
Median	79	10	0.11	20.40	0.25	0.15	23	5	0.25	22.85	0.3	0.21
Maximum	30979	500	53.00	33.70	110.25	78.93	4332	500	104	32.35	62.67	96.27
Minimum	1	1	0.00	10.75	0.00	0.00	1	1	0.00	11.16	0.01	0
Std. Dev.	1264.1	18.16	2.13	2.97	3.87	3.02	157.15	19.4	3.18	3.33	2.09	2.79
Skewness	11	10.93	8.56	0.37	10.86	8.67	7.29	11.02	10.77	0.41	7.99	10.58
Kurtosis	168.44	216.24	121.38	2.74	165.71	124.43	109.26	198.51	210.47	2.65	137.48	203.04
	NP I						NP II					
	Actual Duration (Seconds)	Actual Volume (No of Contracts)	Actual Trading Intensity	Price (In Euros)	Diurnally Adjusted Duration	Diurnally Adjusted Trading Intensity	Actual Duration (Seconds)	Actual Volume (No of Contracts)	Actual Trading Intensity	Price (In Euros)	Diurnally Adjusted Duration	Diurnally Adjusted Trading Intensity
Mean	1912.21	10.81	0.27	20.33	1	1.06	1370.57	10.67	0.27	22.73	0.95	1.01
Median	480	10	0.02	21.20	0.25	0.03	540	10	0.02	22.92	0.38	0.056
Maximum	29933	250	15.00	33.50	17.52	71.92	20542	308	15	37.25	17.54	71.92
Minimum	1	1	0.00	12.00	0	0	1	1	0.00	1.00	0	0
Std. Dev.	3727.12	11.77	1.09	3.00	1.99	3.92	2195.64	14.5	1.09	2.80	1.54	3.92
Skewness	3.71	7.07	7.10	0.64	4.05	8.55	3.28	9.28	7.10	0.35	3.68	8.56
Kurtosis	19.69	92.87	61.60	2.78	23.56	100.11	17.63	130.03	61.60	4.47	23.85	100.11

B.

		Average Duration	Std Duration	Average D/A Duration	Std D/A Duration	Average Volume	Std Volume	Average D/A Trading Intensity	Std D/A Trading Intensity	Average Price	Std Price
						ECX					
2006	Qtr1	5191.40	6801.86	15.73	20.63	17.88	18.38	0.10	0.65	24.58	1.95
	Qtr2	2213.64	3752.81	6.81	11.37	17.91	18.71	0.18	1.04	22.24	3.57
	Qtr3	1847.13	2899.49	5.68	8.85	18.77	27.83	0.20	1.32	17.92	1.26
	Qtr4	619.17	1157.95	1.90	3.60	14.39	21.02	0.56	1.88	17.24	1.26
2007	Qtr1	296.50	515.06	0.92	1.62	12.51	12.59	0.57	1.90	15.10	1.28
	Qtr2	198.31	423.14	0.62	1.41	11.78	14.35	0.77	2.87	21.04	2.53
	Qtr3	173.03	299.37	0.54	0.95	12.93	22.09	0.83	2.71	20.22	0.95
	Qtr4	236.20	507.39	0.73	1.70	10.10	16.42	0.78	2.20	22.54	0.65
2008	Qtr1	82.27	153.79	1.09	2.04	8.81	15.81	0.92	2.67	21.27	1.22
	Qtr2	85.91	164.80	1.14	2.14	10.91	20.13	1.09	3.28	25.73	1.51
	Qtr3	70.98	164.09	0.94	2.22	9.29	17.51	1.22	3.50	24.52	1.90
	Qtr4	66.05	143.73	0.88	1.92	9.86	23.29	1.28	5.09	18.83	2.71
						NP					
2006	Qtr1	15413.75	14195.22	7.59	6.57	13.75	11.09	0.01	0.01	24.08	0.10
	Qtr2	14033.27	9752.76	7.93	5.80	20.78	15.27	0.01	0.02	22.29	4.68
	Qtr3	10113.75	8276.41	5.27	4.58	24.33	25.02	0.02	0.05	17.69	1.34
	Qtr4	5171.03	6330.17	2.80	3.58	15.95	22.14	0.03	0.07	17.57	1.11
2007	Qtr1	2396.67	3953.52	1.25	2.07	11.16	11.11	0.11	0.90	15.28	1.42
	Qtr2	1165.22	2114.56	0.60	1.06	9.56	7.92	0.22	1.09	21.40	2.52
	Qtr3	1792.24	3303.23	0.92	1.66	10.52	12.16	0.30	1.49	20.32	0.97
	Qtr4	1303.02	2505.22	0.68	1.32	10.54	11.41	0.28	1.20	22.59	0.68
2008	Qtr1	1102.16	1772.28	0.76	1.17	11.32	13.93	0.40	1.64	21.24	1.32
	Qtr2	1536.60	2467.28	1.08	1.77	9.76	12.07	0.37	1.41	25.45	1.48
	Qtr3	1616.53	2509.06	1.14	1.82	10.58	14.65	0.15	0.59	24.36	1.87
	Qtr4	1327.78	2033.75	0.91	1.35	10.82	17.66	0.18	0.93	19.97	2.56

Table 3.1 reports the basic statistics of the variables under examination. The first panel of Table 3.1 reports the mean, the median, the max and min, the standard deviation, the skewness and the kurtosis of the actual duration (in seconds), trading volume (contracts per transaction), trading intensity (volume over duration) and price (in Euros), as well as of the diurnally adjusted duration and trading intensity. This panel is divided into four sections one for each market and phase, namely, ECX I, ECX II, NP I and NP II. The second panel of Table 3.1 presents the average and the standard deviation of the same variables, in both markets per quarter. The first section of that panel refers to ECX, while the second refers to NP.

Figure 3.4: Duration and Trading Intensity, Empirical Distributions

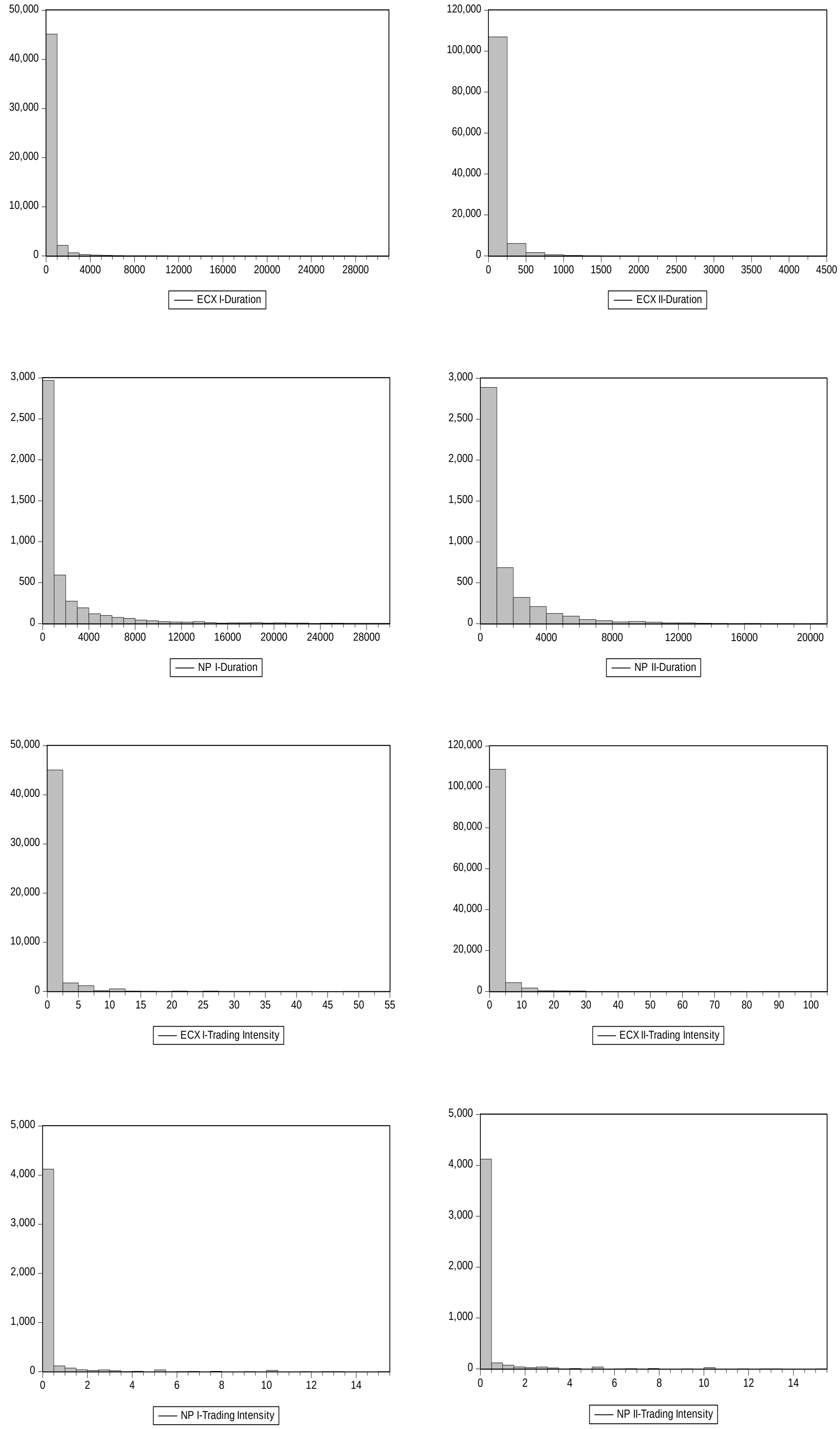


Figure 3.7 presents the distributios of Duration and Trading Intensity across markets and phases

Chapter 4

Duration Modelling in The European Carbon Market

Non-linear Asymmetries and OTC Transactions

4. Duration Modelling in The European Carbon Market

4.1 Introduction

One of the most significant difficulties in modelling intraday data sets is the stochastic nature of the time of occurrence of various events. Unlike daily data, the arrival time of events is irregularly spaced, as a stochastic process. According to the microstructure literature, the arrival time of a transaction might convey price-relevant information, because it measures economic time, which might run faster than calendar time (Engle and Russel, 1998). Aggregation or averaging over fixed time intervals might result in a loss of vital information, therefore a proper modelling of arrival times is required. Duration, defined as the time between two consecutive events, is the variable employed to account for it. Duration appears to be persistent, i.e., it has a memory of a certain length, clustered, and over dispersed, i.e., the standard deviation is larger than the mean. Hence, no concrete conclusions can be drawn using ordinary econometric models, without first adjusting for intraday duration variations.

Engle and Russell (1998), model the persistence and over-dispersion of duration as a dependent point process, using a new model called Autoregressive Conditional Duration (ACD).[91] The modelling of the DGP of duration requires a specification for the conditional mean, conditioned on past durations, and a specification for the density function, which needs to have a positive support.[92] Engle and Russell (1998) employ the linear ARMA specification, as well as the Exponential and the Weibull distributions. ACD bears a very strong resemblance to GARCH models, which has triggered its rapid expansion. Different models have been proposed to deal with different stylized facts and peculiarities of duration series.[93] Several models challenge the simplicity of the conditional mean specification, as inadequate to capture non-linear effects, and distributions of higher complexity are examined for higher accuracy.

A major part of the literature criticizes the inability of the ARMA specification to capture asymmetric effects of past durations. The linear-ARMA has two main drawbacks (Dufour and Engle (2000b). First, there are parameter constraints to ensure

[91] Dependent point process is a stochastic process where each point depends on past values.
[92] Researchers using the ACD framework need to specify an equation for the conditional mean along with a distribution. Standardized durations are explicitly assumed to be iid. Also because of the non-negativity of the durations the potential density functions should have a positive support and thus long right tail. The most widely used distributions belong to the Exponential family. Finally, since the Maximum Likelihood is the method of estimation, the chosen distribution determines the Log-Likelihood function.
[93] For an up-to-date extensive discussion and presentation of various models see Pacurar (2008)

stationarity and avoid negative expectations. Second, it fails to capture non-linear dynamics of higher moments of duration series. The effects of very short or very long durations are assumed to be equal to moderate durations, while another specification, which differentiates between these effects might be more efficient. Dufour and Engle (2000a) postulate that the effect of extreme durations should be different from the impact of “normal” durations, so as to avoid under- or over-prediction.

The first solution proposed by Bauwens and Giot (2000) is the use of the log transformation of the variables. This allows short, lower than the mean, past durations to have a decreasing effect on the expected duration, while long, higher than the mean, durations have an increasing effect. Thus, past durations affect the conditional mean asymmetrically without estimating any additional parameters, or applying any parameter constraints. In Log-ACD there is enhanced sensitivity to clustering, but Dufour and Engle (2000a) criticize the over-adjustment after short durations, due to convergence to minus infinity for very small, close to zero, values of the log function. In contrast, they propose two other models to deal with asymmetric effects of past duration in a data-driven way. First they suggest the Box-Cox-ACD (BCACD), which is a Box-Cox transformation of duration series. This allows the data to determine the size of the effect of past durations. The model nests the linear and the Log-ACD models as special cases. Second, they extend Nelson’s (1991) EGARCH model using a piecewise linear parameterization of the conditional mean specification. Thus, the effect of past durations remains linear, but it also allows past durations that are considerably larger or shorter than the mean to have a different impact on the expected duration.

Other studies employ a regime-switching framework to capture strong asymmetries that cannot be summarized in a single coefficient. Zhang et al. (2001) propose the Threshold-ACD (TACD), which allows for a piecewise modelling of expected duration, where the conditional mean specification in each regime can be linear or non-linear. Meitz and Teräsvirta (2006) extend TACD model (Zhang et al., 2001) to allow for a smooth transition from one regime to another. This model is referred to as the Smooth Transition ACD model (ST-ACD). The regime-switching framework adds flexibility to duration modelling, where other economically relevant variables can have an indirect impact on duration expectations. This idea will be further developed in the next chapter.

Another branch of the literature tests the appropriateness of the distribution. They challenge the simplicity of the Exponential distribution as insufficient for capturing higher moments of duration. The most common distributions used are the Exponential (Engle and Russel, 1998), the Weibull (Engle and Russell, 1998) and the Generalized Gamma (Zhang et al., 2001). These three distributions are nested, allowing for generalization and flexibility. Grammig and Maurer (2000) propose Burr as a mixture of Weibull distributions, which includes the Exponential, the Weibull and the Log-logistic distributions as special cases; Hautsch (2002) employs the Generalized F distribution; De Luca and Zuccolotto (2006) utilize the Pareto type-II distribution. These studies provide evidence that the fitting and forecasting power of the models varies depending on the data set employed and on the individual market characteristics. Usually, higher flexibility is rewarded with better fitting and higher forecasting accuracy (Dufour and Engle, 2000a; Bauwens et al., 2004; De Luca and Zuccolotto, 2006).

Furthermore, Bauwens et al. (2004) postulate that durations exhibit an idiosyncrasy that cannot be captured by a single density function. The implication of this statement is twofold. First, according to Grammig and Wellner (2002), the conditional mean specifications should be of utmost importance compared to the density functions. They argue that the marginal contribution of higher complexity in the assumed distributions is limited and that usually fails to capture market stylized facts. Consequently, a misspecification of the conditional mean would have a greater impact on expectations, compared to a misspecification in the assumed distribution. Second, other studies use a different statistical approach that allows for a mixture of different distributions to describe the DGP of durations. The different regimes of durations, which are associated with different density functions, can be determined by either past durations (i.e., self-exciting process) or another economically relevant variable. De Luca and Gallo (2004) and De Luca and Zuccolotto (2006) propose a discrete mixture of distributions, while Bauwens and Veredas (2004) and Ghysels et al. (2004) propose a continuous mixture of distributions. In parallel, Hujer et al. (2002) and Hujer and Vuletic (2007) propose a Markov-Switching-ACD model where the threshold variable is latent.

The new dynamics of the mixture of distributions have implications for various microstructure issues that exceed the scope of the current chapter. The regime-switching framework allows the connection of each regime with special features of the market, opening new horizons in market microstructure analysis. In addition, these models allow other variables to affect duration expectations, thus increasing significantly flexibility when examining inter-relations between market variables. Several issues, such as information dissemination, liquidity and pricing, have been debated in the literature and need to be developed further. These topics constitute the main analytical focus of the following empirical chapters. In more detail, Chapter 5 presents an analysis of the relation between duration, trading intensity and information, while Chapter 6 focuses on their price effects. The current chapter provides the foundation for this analysis.

The vast majority of the afore-mentioned studies employs data sets from liquid and well-established markets. The Carbon market, however, appears to have some unique characteristics that might challenge the suitability of some ACD models. The market is relatively new and relatively less liquid. But it is progressively gaining complexity and "depth", while remaining politically influenced. This creates a highly speculative environment where liquidity, information asymmetry and regulation play an important part (Viswanathan, 2010). This market might exhibit different trading patterns driven by different stylized facts, and previously applied models may not be appropriate without further empirical adjustments.

One of the most important facts is the presence of OTC transactions. OTC EUA holders are allowed to enter the market directly in Nord Pool and use Exchange For Swap (EFS) or Exchange For Physical (EFP) facilities in ECX. Considering that the non-organized market existed before the organized, and that OTC transactions are larger than the organized exchange transactions, they might decide to register their positions for a reason.[94] Therefore, these trades might convey price-related information that could result in more frequent trading, and therefore shorter durations. In contrast, given the "cap and trade" character of the market, these trades might express excessive liquidity needs.[95] Given their large size in combination with the relatively low market liquidity, their presence might increase the expected duration, either because they deter other traders from entering the market or simply because they consume current liquidity. Furthermore, OTC transactions account for about 30 percent of the total transactions in ECX and for about half of the transactions in Nord Pool. These transactions might have a different impact on expected durations that cannot be ignored.

[94] Indeed, countries other than those in ANNEX I (i.e., countries that participate in emissions reduction scheme under the Kyoto Protocol) can enter the market, and their motivation for trade is not only compliance. (see, inter alia, "State and Trends of the Carbon Market", Information Emissions Trading Association (IETA), 2009).

[95] In the literature two types of traders are recognized, informed and uninformed traders. Informed traders are the investors who have private information and trade to make money by using it. On the contrary, uninformed traders transact because of various other reasons, like portfolio optimization, risk management or because they simply need or want to. This is particularly relevant in the EU ETS, since companies need to be able to present each April the allowances related to their energy consumption. These traders are often called liquidity traders due to the fact that they supply liquidity to the market.

Several comments related to market structure can also be made. First, the legislation applied in each phase of the market is different and it is in line with the predetermined aims of emission reductions for each period. Consequently, different trading patterns might dominate each phase and should be examined separately. Moreover, ECX and Nord Pool are two different exchanges. Although the same asset is traded in both, there are significant differences. ECX is far more liquid, offers more instruments and differs in structure. Geographical location and investment environments are important aspects as well, since ECX is based in London while NP in Norway.[96] Thus each market and each phase should be treated separately, allowing for different parameterizations. EU ETS market is a new market and some financial instruments such as EUA futures have been introduced later on. Their trading has gradually gained complexity, and durations appear to be larger at the beginning of trading. Similarly, the fixed maturity date of these contracts is expected to seasonally influence the realized durations. "Long-life" future contracts are expected to exhibit fluctuations in the number and frequency of trading throughout their trading time, depending on various factors, such as seasonal patterns or OTC transactions.

Another influential fact is trading intensity. In a relatively less liquid market, such as the EU ETS, liquidity is expected to play an important role in determining trading conditions. The available liquidity is undeniably a main determinant of the ability of market participants to develop their trading strategies, as well as of how information is resolved into prices. Higher trading intensity might help in incorporating new information faster, but according to Grammig et al. (2007) it might provide coverage for informed traders. Depending on market structure and sensitivity of market participants in anticipating incoming information, in the first case trading is expected to be more intensive because information is quickly revealed, while in the second case, trading should not be affected since informed traders have more time to exploit their information advantage, due to the fact that they can conceal their actions for longer. The analysis in this present and the following two chapters is supportive of the first scenario that information dissemination becomes faster as trading size and frequency increase.

EU ETS has become the major mechanism for emission reduction, but the microstructure of the market has yet to be examined in depth. Analytical efforts focus mainly on price dynamics and the unique features of the underlying commodity. It

[96] For a further discussion see "Derivative Instruments in the EU ETS an early market Perspective" Uhrig-Homburg and Wagner (2007).

appears that there is no relevant study that explicitly models time between events. The rapid expansion of the market, its increasing complexity, and its political influence might create a significant asymmetric information dynamics over time and across markets. Different phases and structures might attract different traders with different needs. Their trading strategies are realized in transactions instigated at specific points in time and, therefore, patterns in durations might express the combination of different trading groups. These groups are expected to have an asymmetric impact on expectations. In addition, abnormalities and extreme values are expected, especially in early stages, and more flexible distributions might better describe the DGP of duration. Finally, the potentially increasing impact of OTC transactions is also examined.

The present chapter constitutes primarily an attempt to provide some evidence on duration modelling in the EUA futures market considering various stylized facts. This contributes to the literature in various ways. First, an extensive examination of trading patterns in the European Carbon market is conducted, focusing on the time dimension of trading. This provides a solid ground for further analysis concerning the information that can be extracted from the arrival times of various events. Second, several stylized facts, such as increasing market liquidity and OTC transactions, are incorporated in the analysis. This further examines the trading environment of the market, and empirically extends ACD literature. ACD specifications that have been employed in more advanced markets might not hold in such a unique market, and, according to Bauwens et al. (2004), a more explicit conditional mean specification might significantly improve model performance. Third, the models proposed in this analysis, further extend the ACD literature, concerning the asymmetric effect of extreme durations. Two new models are proposed, whereby past durations are allowed to have an asymmetrically non-linear effect on expectations, which depends on other, economically relevant variables. Thus, exogenous variables are allowed to have an impact on the conditional mean without challenging the exogeneity assumption (Engle and Russel, 1998). In this analysis, the impact of trading intensity and OTC transactions is examined, but this framework can be extended to account for other factors.

In more detail, the fitting and the forecasting ability of various existing models are initially examined. The original ACD, the Log-ACD, the BCACD and the EX-ACD have been chosen, along with the Exponential, the Weibull and the Generalized Gamma distributions to examine whether higher complexity is needed. Furthermore, two new models are proposed to account for non-linear asymmetries of trading intensity and

OTC transactions. Drawing on the BCACD model, the size of the effect of past durations is allowed to smoothly vary (Meitz and Teräsvirta, 2006) across another variable and is data-driven. The new model practically introduces a highly non-linear component in duration modelling, which is associated with duration and trading intensity. Another specification that allows for the impact of past durations to be revised according to another variable is also proposed. The effect of past durations is still linear on every regime of the variable employed, but the overall effect is non-linear.

The most important finding of the empirical analysis presented in the chapter maintains that empirical adjustments of the ACD specifications, significantly improve the performance of the models in the European Carbon market. This confirms the claim by Bauwens et al. (2004) that the correct specification of the mean equation contributes more to duration modelling than the distributional assumptions. However, the right choice of an appropriate density function further increases accuracy. In more detail, the linear models consistently outperform their non-linear counterparts, while the Generalized Gamma distribution outperforms the other two distributions. It is also found that OTC transactions have an increasing effect on expected duration. This can be explained either by information or liquidity effects. The analyses, in Chapters 5 and 6, will show that it is more likely to be related to information.

Furthermore, past durations and trading intensity provide consistent results, indicating that intense trading is followed by more intense trading, while in relatively less active stages of the market, longer durations are expected. Both transaction size and trading frequency appear to have a long-lasting effect and higher volume of trading decreases expected duration. This could be an indication of episodes of intense trading, which according to Easley and O’Hara (1992), could be a sign of increased presence of information. Building on this idea and taking into account the highly speculative nature of the Carbon market, these intense trading episodes may indicate episodes of exogenous information, which is gradually being revealed through the trading process.

In addition, OTC models improve model performance, especially in Nord Pool and in Phase II. This could indicate increased informational content of OTC. Other market participants appear to react more quickly to exogenous information than to order flow signals. Trading intensity attempts to capture these signals. Therefore, although large or fast transactions tend to increase the frequency of trading, an OTC transaction, whenever it occurs, seems to deter trading, probably due to asymmetric information. In contrast, ECX appears to be more affected by order flow imbalances, which indicates higher complexity and a more mature trading environment.

These findings are particularly relevant to regulatory authorities, as a better understanding of intraday trading activity can enhance market regulation by finding a balance between market innovation and liquidity needs. According to Viswanathan (2010), this would provide the foundations of more accurate pricing and would help the market serve its purpose of reducing emissions. A more precise duration model could also improve market making practices in EU ETS. Market makers can better manage risk if they know how long (i.e., expected duration) they are going to be exposed to it, and this would result in narrower spreads. Limit order traders can also benefit from a better duration model, since they can develop better trading strategies by managing the time dimension of "execution" risk.

The next section, 4.2, discusses the methodology employed, while section 4.3 presents the empirical results. The last section, 4.4, summarizes the main findings.

4.2 Methodology

4.2.1 Previous studies

The Basic ACD Model

Engle and Russell (1998) propose the ACD models for high frequency irregularly spaced data. They model the Data Generation Process of inter-trade intervals, namely durations, x_i, as a dependent stochastic process. ACD models are a class of dependent point processes (i.e., stochastic processes that generate sequences of time intervals). The conditional mean, $E(x_i|x_{i-1}, \dots, x_1)$, of duration, x_i, varies over time as a function of past durations.[97] ACD is formalised as:

$$x_i = \psi_i \varepsilon_i \tag{4.1}$$

$$\psi_i = \psi(x_{i-1}, \dots, x_1; \varphi_1) \tag{4.2}$$

$$\varepsilon_i \sim i.i.d. with\ density\ f(\varepsilon_i; \varphi_2) \tag{4.3}$$

[97] The ACD framework allows for the inclusion of other economically relevant factors in the conditional mean specification. However, the vast majority of studies assumes exogeneity and allows durations to be determined only by past durations. In this study, only linear and non-linear conditional mean specifications that condition duration on its past values will be employed, for comparability reasons.

where x_i is the diurnally adjusted duration, ψ_i is the expected duration conditioned on past durations, ε_i are the residuals, otherwise called the standardized durations, and φ_1 and φ_2 are vectors of parameters, stacked in the vector $\theta = (\varphi'_1\ \varphi'_2)$.

This general model allows researchers to choose various parameterizations and formulations for the conditional mean, as well as for the density functions, ε_i. Different formulations of the conditional mean equation can be employed as long as they model the expected duration on past values of x_i and a function of past durations, $g(x_i)$, or a function of past standardized durations, $p(\varepsilon_i)$.[98] In addition, any distribution defined on a positive support can be specified as a density function $f(\varepsilon_i)$; for different distributions see Lancaster (1997). The most popular density functions used in the literature are the ones belonging to the Exponential family of distributions.[99] This way, various combinations are made available and researchers have the freedom to model durations using the parameterizations and the distributions that best fit the data set employed. The function in Eq. (4.2) can either be linear or non-linear, perhaps drawing on the stylized facts of the market examined. Moreover, life distributions, which have positive support and long right tales, with more parameters can be chosen for a better fitting. However, usually greater flexibility comes at a cost of higher complexity. The various specifications considered in this analysis are presented next.

Conditional Mean Equation Specifications

ARMA-specification

The simplest version, proposed by Engle and Russell (1998), is the linear-ARMA specification (ACD):

$$\psi_i = \omega + \sum_{j=0}^{m} a_j x_{i-j} + \sum_{j=0}^{q} \beta_j \psi_{i-j}, \tag{4.4}$$

where ω (*omega*), α (*alpha*) and β (*beta*) are parameters to be estimated, ψ is the expected duration and *x* is the realized duration. The expected duration is modelled as a linear combination of past actual and expected durations.

[98] In the ACD literature, in many parameterizations of the conditional mean equation the standardized durations $\varepsilon_i = x_i/\psi_i$ are used instead of the normal durations.
[99] Engle and Russell (1998) use Exponential and Weibull distributions. Some other studies also use a mixture of distributions belonging to the Exponential family. Grammig and Maurer (2000) use a mixture of Weibull distributions to estimate the so-called Burr-ACD model.

Log-ACD

Another specification that introduces asymmetry between the effects of long and short past durations on expected duration is the Logarithmic ACD (Log-ACD) (Bauwens and Giot, 2000).[100]

$$ln\,\psi_i = \omega + \sum_{j=0}^{m} a_j ln\, x_{i-j} + \sum_{j=0}^{q} \beta_j ln\,\psi_{i-j}, \tag{4.5}$$

or

$$ln\,\psi_i = \omega + \sum_{j=0}^{m} a_j ln\, \varepsilon_{i-j} + \sum_{j=0}^{q} \beta'_j ln\,\psi_{i-j}. \tag{4.6}$$

In this specification, for positive *a*'s the effect of durations larger than the mean ($x_i > 1\ or\ \varepsilon_i > 1$) are separated from the effect of durations shorter than the mean ($x_i < 1\ or\ \varepsilon_i < 1$). Durations, larger than the mean, have a positive and rather increasing effect, while shorter durations have a negative effect.

EXACD

Dufour and Engle (2000a) propose a piecewise linear specification (Exponential ACD) (EXACD).

$$ln\,\psi_i = \omega + \sum_{j=0}^{m}\left[a_j \varepsilon_{i-j} + \zeta_i\left|\varepsilon_{i-j} - 1\right|\right] + \sum_{j=0}^{q} \beta_j ln\,\psi_{i-j}, \tag{4.7}$$

where ζ (*zeta*) is a parameter to be estimated that captures the effect of the dispersion of the mean.[101] This is a peace-wise linear specification. The effects of past durations are still linear. However, when durations are shorter than the mean ($\varepsilon_i < 1$), the slope of their effect on the conditional mean is $(\alpha - \zeta)$ with an intercept $(\omega + \zeta)$ and when durations are longer than the mean ($\varepsilon_i < 1$), the slope is $(\alpha + \zeta)$ with an intercept $(\omega - \zeta)$. In this model, the distance from the mean, where $\left|\varepsilon_{i-j} - 1\right|$ measures how far the realized duration is from the expected, acts as a correction factor to the *ω* and the

[100] Where $\varepsilon_i = \frac{x_i}{\psi_{i+1}}$ and is the standardized duration and $\beta'_j = \beta_j + a_j$. In addition, for a comparison see Fernandes and Grammig (2006). Readers familiarised with GARCH literature, can easily spot the similarities with Log-GARCH.

[101] Where $\varepsilon_i = \frac{x_i}{\psi_{i+1}}$ is the standardized duration. Readers familiar with GARCH literature can easily spot the similarities with EGARCH.

α´s estimated. The effect of past realized durations that are close to the expected, and thus $\varepsilon_i \approx 1$, is different from the "abnormal" ones that are not appropriately predicted. When $\zeta > 0$, the expected duration increases when abnormal returns are observed, either very short or very long durations. In contrast, the expected duration decreases when $\zeta > 0$. This effect is stronger the farther is the realized duration from the expected.[102]

BCACD

Dufour and Engle (2000a) also propose another specification. The so-called Box-Cox ACD (BCACD) is given by:

$$ln\,\psi_i = \omega + \sum_{j=0}^{m} a_j(\varepsilon_{i-j})^{\delta} + \sum_{j=0}^{q} \beta_j ln\,\psi_{i-j} \qquad (4.8)$$

where δ (*delta*) is a parameter to be estimated indicating the size of the effect of the realized durations.[103] This model nests the first two specifications as special cases. The linear when $\delta = 1$, and the logarithmic when $\delta \rightarrow 1$. Here the impact of the shocks (past durations) on the conditional mean is data-driven.

Conditional Density Function Specifications

According to the above, the specification of an ACD model follows two steps. The researcher needs to identify a conditional mean equation and then a density function for the residuals. The density functions that can be employed need to have a positive support and the most popular in the literature belong to the Exponential family. Following Dufour and Engle (2000a) the following distributions will be employed in the analysis of this chapter.

Exponential

$$f(\varepsilon_i \mid x_{i-1}, \dots, x_1; \theta) \equiv {}^{1}\!/_{\psi_i} \exp\left(-{}^{x_i}\!/_{\psi_i}\right) \qquad (4.9)$$

[102] The same authors propose a more general form, where the threshold does not necessarily need to be one and the model can be generalized with a number of thresholds K>1. More specifically, the generalization of EXACD, for thresholds K>1 could be:
$ln\,\psi_i = \omega + \sum_{j=0}^{m}\left[\alpha_{i-j}\varepsilon_{i-j} + \sum_{k=1}^{K}\left|\varepsilon_{i-j} - \tau_k\right|\right] + \sum_{j=0}^{q}\beta_j\, ln\,\psi_{i-j}$

[103] Where $\varepsilon_i = \frac{x_i}{\psi_{i+1}}$ is the standardized duration. Readers familiarised with GARCH literature, can easily spot the similarities with Power-GARCH (PGARCH).

Weibull

$$f(\varepsilon_i \mid x_{i-1}, \dots, x_1; \theta) \equiv {}^{\gamma}\!/_{\chi_i} \left[\frac{x_i \Gamma(1 + {}^{1}\!/_{\gamma})}{\psi_i}\right]^{\gamma} \exp\left(-\left[\frac{x_i \Gamma(1 + {}^{1}\!/_{\gamma})}{\psi_i}\right]^{\gamma}\right) \tag{4.10}$$

Generalized Gamma

$$f(\varepsilon_i \mid x_{i-1}, \dots, x_1; \theta) \equiv \frac{\gamma}{\chi_i \Gamma(\lambda)} \left[\frac{x_i \Gamma(\lambda + {}^{1}\!/_{\gamma})}{\psi_i}\right]^{\gamma\lambda} \exp\left(-\left[\frac{x_i \Gamma(\lambda + {}^{1}\!/_{\gamma})}{\psi_i \Gamma(\lambda)}\right]^{\gamma}\right) \tag{4.11}$$

In these equations $\Gamma(\cdot)$ is the gamma function of the distribution parameters $\varphi_2 = \gamma, \lambda$, where $\gamma > 0$ and $\lambda > 0$.

These three distributions belong to the Exponential family and are nested. When $\lambda = 1$ the Generalized Gamma nests the Weibull distribution, and when $\gamma = 1$ the Weibull nests the Exponential distribution. More specifically, the Exponential distribution is a decreasing function while the Weibull assigns higher (lower) probability to extreme values (i.e., very short or very long durations) when $\gamma < 1$ $(\gamma > 1)$. The Generalized Gamma distribution allows γ to vary and offers higher flexibility.

However, what is of particular interest in survival analysis is the hazard function.[104] The hazard function is defined as the ratio of the probability of the density function over the survival function.[105] This is an instantaneous rate of transition from a non-transaction state to a transaction state. In other words, hazard functions measure the probability of a transaction to occur at a particular point in time, given that no transaction has occurred till now. This is different at every point of time and consequently different data sets will have different hazard functions. In addition, different distributions, depending on their parameters, have different shaped hazard functions. In particular, the theory of Survival analysis holds that the hazard function of the Exponential distribution is constant. In

[104] Two remarks should be mentioned here concerning the use of the term “the hazard function”. First, the term hazard function is associated with cross sectional data, while the term “intensity” is used more often in time series. In addition, since the distribution functions refer to the conditional density of durations, it is more appropriate to use the term “conditional intensity”. However, mainly for consistency reasons with the literature the term hazard will be used throughout this thesis. This will also avoid confusing the hazard function with “trading intensity”, which is defined as the ratio of volume over duration as in Eq. (4.16).

[105] Survival function is defined as the probability of an event to occur after a predetermined time Y. For a more extensive discussion see Peter J. Smith “*Analysis of failure and survival data*” Chapman & Hall/CRC (2002).

contrast, the hazard function of the Weibull distribution is monotonically decreasing (increasing) when $\gamma < 1$ ($\gamma > 1$). The economic meaning of a decreasing hazard is that the probability of occurrence of a transaction decreases with time. In the more flexible Generalized Gamma distribution, the hazard function is non-monotonic (Lunde, 2000). When $\lambda\gamma > 1$ and $\gamma < 1$, the hazard function follows an inverted U-shape pattern, indicating that the probability of a transaction to occur now increases up to a specific duration (i.e., time since last transaction) and then decreases. In the opposite case of $\lambda\gamma < 1$, and $\gamma > 1$ the hazard function follows a normal U-shape pattern.

4.2.2 Extensions proposed

ACD-OTC

One of the main questions raised in this study is whether OTC transactions follow a different trading pattern than normal trades and how that affects expected duration. Statistical analysis of the data indicates a large amount of OTC transactions in both exchanges. Almost half of the transactions in Nord Pool are made by OTC EUA holders. In addition, the size of some of these transactions exceeds the average transaction size of normal transactions, sometimes even by hundred times. Large volume trades can be an indication of new information or significant changes in fundamental market forces that have yet to reach the organized market. Furthermore, sometimes these trades seem to come in blocks of transactions.[106]

In order to identify any potential effects of OTC transactions the following empirical extension of the basic ACD model is employed. Starting from the linear ACD model of Engle and Russell (1998), a dummy variable is introduced to account for any different effects that OTC transaction might have.

The new model is formulated as:

[106] This is particularly apparent in Nord Pool. A close examination of the data set provides evidence of OTC transactions coming in waves. This could be anticipated as an information inflow signal that might be followed by participants of the organized market. On the contrary, it could be seen as a signal from the market, where information is exported and OTC EUA holders enter the market to take advantage of it. Finally, it could simply be a lack of demand or supply, where OTC EUA holders want to enter the market with a large volume, but there are no sufficient counter orders.
This could probably be investigated in greater depth when other variables, such as trade direction or price movements, are also taken into account. However, following the relevant literature, this study uses the ACD framework to model duration without using any other marks. Therefore, in order to be consistent and have comparable results, only past durations are allowed to have an impact on duration. Even in the cases of ACD-OTC or switching regimes models, duration is modelled on past durations, organized though in groups determined by OTC transactions or trading intensity.

$$\psi_i = \omega + \sum_{j=0}^{m} (a_j + \zeta_j * D_{i-j}) x_{i-j} + \sum_{j=0}^{q} \beta_j \psi_{i-j}, \tag{4.12}$$

where ζ (*zeta*) is the parameter that captures the effect of OTC transactions and

$$D = \begin{cases} 0 & for\ non - OTC\ transactions \\ 1 & for\ OTC\ transactions. \end{cases} \tag{4.13}$$

In this model, OTC transactions are allowed to have an impact on the coefficient of realized durations, α. In this case, α, compared to the linear model, represents the average impact of past durations on the expected duration. *Zeta*, ζ, is allowed to change this average impact when the last transaction, or any other past transaction depending on the memory length of the model, is an OTC transaction. When ζ is statistically significant, OTC transactions appear to have a significant impact on the DGP of durations, which means that OTC transactions probably follow different patterns from normal market transactions.

The sign of ζ allows for a further interpretation. When $\zeta > 0$ ($\zeta < 0$), it means that, when the last transaction is an OTC transaction, the expected duration is longer (shorter). Various economical explanations might link the sign of ζ to an economic event. First, a positive ζ could possibly mean that, when an OTC transaction is reported to the market, market members understand its presence as information inflow and they seem reluctant to trade for the fear of losing money by trading with informed traders. Second, a positive ζ could possibly mean that these transactions "exhaust" the current demand or supply, since some OTC transactions are really large, and the market needs some time to find a new balance, thus the longer durations. In contrast, a negative ζ could indicate an inflow of information but for a different reason. Traders obsever OTC transactions as informative and try to follow their trading pattern. Whatever the sign is, its interpretation lies in the microstructure of the market under examination.

In order to take into account any non-linear effect, reported extensively in the ACD literature, the BCACD parameterization is also employed. The new model, BCACD-OTC, could be formulated as:

$$ln\, \psi_i = \omega + \sum_{j=0}^{m} (a_j + \zeta_j * D_{i-j}) * (\varepsilon_{i-j})^{\delta} + \sum_{j=0}^{q} \beta_j ln\, \psi_{i-j}, \tag{4.14}$$

where notation is defined as above. Note that the coefficient *delta* (δ) is also influenced by the insertion of the OTC dummy variable. It still remains to be data-driven, but in this case it measures the average size of the impact of past durations. Indeed, the total coefficient of past durations (i.e., $\alpha + \zeta$) is not constant, depending on whether the last transaction was an OTC one.[107]

ST-BCACD

A significant issue that has been discussed extensively in the ACD literature is the asymmetric effects of past durations on expected duration. Many studies report that non-linearity issues arise when ACD models are employed. Two of the most widely used methods are the non-linear and the threshold specifications. In this study a combination is employed by allowing the size of the impact of past durations on the expected duration to vary across different regimes. A general title for the models proposed could be Smooth Transition BCACD models.[108]

More specifically, features of the non-linear BCACD model of Dufour and Engle (2000a) are combined with the Smooth Transition-ACD model of Meitz and Teräsvirta (2006). The main idea is that the impact of past durations on the conditional mean could depend on other economically relevant factors as well. The values that these factors can take on can be divided into regimes that may possibly determine different effects of past durations on the conditional mean.[109] These different effects might indicate asymmetric effects of past durations.

Therefore, the new model proposed is formulated as:

---

[107] Delta can be allowed to differ between normal and OTC transactions. In that case the model could be formulated as:

$$ln\psi_i = \omega + \sum_{j=0}^{m} a_j(\varepsilon_{i-j})^{\delta 1} + (\zeta * D_{i-j}) * (\varepsilon_{i-j})^{\delta 2} + \sum_{j=0}^{q} \beta_j ln\, \psi_{i-j} \ ,$$

where δ_1 is the size of the effect of the duration of normal transactions, while $\delta 2$ is the size of the effect of the durations of the OTC transactions. In the BCACD-OTC model though delta is common for both effects (it is chosen this way for parsimony and comparison reasons) and it works as an average.
Furthermore, if the size of the impact of past durations is of particular interest, a probably more effective linear model would be:

$$\psi_i = \omega + \sum_{j=0}^{m} a_j x_{i-j} + (\zeta * D_{i-j}) x_{i-j}^{\delta} + \sum_{j=0}^{q} \beta_j \psi_{i-j} \ ,$$

where delta in that case captures the size of the effect of OTC transactions only.

[108] Several issues can be explored in dealing with seasonalities other than intraday seasonalities. Seasonal and market dummies can be employed. Data sets can be separated into segments belonging to a different phase or a different market, but this is like working with dummy variables. However, this way the long memory of observed durations is being ignored. To deal with it, instead of using additional dummy variables, the idea of regime-switching models is used here.

[109] In order for these economically relevant factors to have economic meaning ,they need to be derived by values observable variables, such as price, volume etc. that come along with time stamps. In the literature these events are called marks.

$$ln\,\psi_i = \omega + \sum_{j=0}^{m} a_j(\varepsilon_{i-j})^{\delta'} + \sum_{j=0}^{q} \beta_j ln\,\psi_{i-j}, \tag{4.15}$$

where notation is as above, apart from δ, which is allowed to vary across a threshold variable. The regimes of this threshold variable determine different δ coefficients and, therefore, different sizes of the effect of past durations. The threshold variables that appear to be more related to durations are past durations and trading intensity, which is used to proxy the state of the market.

Meitz and Teräsvirta (2006) are the first to implement ACD models with the STAR framework, as a self-exciting process. The initial idea of BCACD models is that short durations affect the expected duration differently than long durations do. The difference should be even larger after very short or very long durations.

In addition, investors are believed to gain information from various transaction-related sources. One of them is how "thick" (i.e., how active) the market is at any particular moment. This "knowledge" allows them to have an opinion about the "fair" value of the underlying asset at the time, and it influences when and how they transact. Consequently, along with past durations, some other variables might affect the expected duration. One variable, which is also present in the literature, is lagged trading intensity. In this study trading intensity is used to capture the "thickness" of the market in the context of trading activity. Trading intensity of transaction $i-j$, S_i, is defined as:

$$S_i = \frac{v_{i-j}}{d_{i-j}} \tag{4.16}$$

where t_i is the time stamp of trade i, $d_i = t_i - t_{i-1}$is raw duration and v_i is volume or transaction size, measured as number of contracts, of trade i. When high trading volume and/or high trading frequency, measured as short duration, are observed S_i increases and this is interpreted as a "thick" or active stage of the market. j is restricted to one, allowing only the last transaction to have an impact on the size of the effect of past durations, and thus an asymmetric effect on the conditional mean.

Two stages of this threshold variable (i.e., two regimes of trading intensity) will be used. This allows for potentially different effects of past durations. When the duration of the previous transactions is higher (lower) than the threshold value, the previous value of trading intensity indicates a "thin" ("thick") market. Moreover, since a gradual adjustment of the expected duration to the changing market conditions could be more relevant, a smooth transition between the regimes is assumed. Formally, δ' is defined as:

$$\delta' = \delta_1 * \left(1 - G(S_i: g, s)\right) + \delta_2 * G(S_i: g, s), \tag{4.17}$$

where δ_1 and δ_2 are coefficient parameters, to be estimated, that measure the size of the effect of past durations in regime one and two respectively; $G(S_i: g, s)$, is the smooth transition function used, where g is the smoothness parameter, where smaller values indicate smoother the transition, S_i is the economically relevant variable chosen as the threshold variable, which is specified to be trading intensity, and s is the threshold value of the threshold variable that determines the regimes. Similar to Meitz and Teräsvirta (2006), the general form of smooth transition function used is the logarithmic function, which is written as:

$$G(S_i: g, s) = \left(1 + exp\left\{-g \prod_{k=1}^{K} (S_i - s_k)\right\}\right)^{-1} \tag{4.18}$$

where $g > 0$ and k is restricted to 1, indicating high and low states of the threshold variable. This way, for every duration, δ' is a weighted average of the coefficients δ_1 and δ_2 (i.e., the coefficients of regime one and two respectively. When the threshold variable is considerably smaller than the threshold value, $S_i \ll s_k$, the transition function is tends to one, $G(S_i: g, s) \to 1$, and therefore the δ' tends to δ_1, $\delta' \to \delta_1$. In contrast, when $S_i \gg s_k$, $G(S_i: g, s) \to 0$ and $\delta' \to \delta_2$.

Finally, just for clarity reasons, when past durations are employed as the threshold variable, the model is called Self-Exciting Smooth Transition BCACD (SEST-BCACD). But, when trading intensity is employed, the model is abbreviated as STV-BCACD Smooth Transition Volume BCACD (STV-BCACD) to indicate the implementation of volume into the analysis.

4.2.3 Estimation

The estimation method used is maximum likelihood using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) (1970) algorithm with numerical derivatives. The Log-Likelihood functions, $L(\Theta)$, maximized, depending on the error distribution assumed, are as follows.

For the Exponential (henceforth abbreviated E) distribution:[110]

$$L(\Theta) = -\sum_{i=1}^{N(T)} \left\{ \log(\psi_i) + \frac{\chi_i}{\psi_i} \right\} \tag{4.19}$$

For the Weibull (henceforth abbreviated W) distribution:

$$L(\Theta) = \sum_{i=1}^{N(T)} \ln\left(\frac{\gamma}{\chi_i}\right) + \gamma ln\left(\frac{\Gamma\left(1 + {}^{1}/_{\gamma}\right)\chi_i}{\psi_i}\right) - \left(\frac{\Gamma\left(1 + {}^{1}/_{\gamma}\right)\chi_i}{\psi_i}\right)^{\gamma} \tag{4.20}$$

For the Generalized Gamma (henceforth abbreviated G) distribution:

$$L(\Theta) = \sum_{i=1}^{N(T)} \ln\left(\frac{\gamma}{\chi_i}\right) - \ln\left(\Gamma(\lambda)\right) + \gamma\lambda ln\left(\frac{\Gamma\left(\lambda + {}^{1}/_{\gamma}\right)\chi_i}{\psi_i \Gamma(\lambda)}\right) - \left(\frac{\Gamma\left(\lambda + {}^{1}/_{\gamma}\right)\chi_i}{\psi_i \Gamma(\lambda)}\right)^{\gamma} \tag{4.21}$$

4.2.4 Performance

In-sample Modelling

Following the literature, the first test will be the remaining autocorrelation in the squared standardized durations, $\left({}^{\chi_t}/_{\psi_t}\right)^2$.[111] Under the null hypothesis of no autocorrelation left, the Ljung-Box statistics (i.e., Q-statistics) at 15 lags, which is the lag length used in the literature, should be lower than 25.[112] In addition, the Schwarz/Bayesian information criterion is calculated.[113] Wald hypothesis tests for the additional conditional mean specification and the distribution parameters will be calculated.

---

[110] Engle and Russell (1998) propose this Log-Likelihood function for exponential distribution, showing that the maximizer of L(Θ) will be consistent and asymptotically normal with a covariance matrix given by the familiar robust standard errors from Lee Hansen.

[111] The rationale for using the remaining autocorrelation as a measure of goodness of fit originates back to the Engle and Russell (1998) study.

[112] The test statistic is distributed as a χ^2_{15} with a 5 percent critical value of 25.

[113] (BIC = [-2*(Log-Likelihood) + k * ln(R)]/R). The lower the BIC is the better. BIC penalizes over parameterized models. k is the total number of estimated parameters (model and distribution). R stands for the number of observations. BIC is more conservative than Akaike information criterion (AIC) and is usually preferred.

For the Weibull distribution the null hypothesis is that $\gamma = 1$, which is the case when the distribution reduces to Exponential. For the Generalized Gamma there are three null hypotheses, γ=1, λ=1 and γ=λ=1. When λ=1 G reduces to the W; when γ and λ are jointly one G reduces to the E. Furthermore, in the EXACD model the null hypothesis is that ζ=0, where no asymmetric effect, measured as the distance from the mean, is present. In BCACD the null hypothesis is that δ=1, where the size of the impact of past durations is determined only by the parameter α. In the ACD-OTC framework, the null hypothesis for ζ is that it is zero. When it cannot be rejected no OTC effects can be assumed. Moreover, the null hypothesis for δ is similar to the BCACD model.

Finally, the in-sample long- and short-term "forecasts" will be used as an additional measure of goodness of fit, along with the *KS-statistics.*[114]

Out-of-sample Forecasting Analysis

Out-of-sample one-step forecasts can be easily generated using the conditional mean Eqs. (4.4), (4.5), (4.7), (4.8), (4.12), (4.14) and (4.15). However, for generating the ten-step forecasts the methodology of Dufour and Engle (2000a) will be employed.[115] More specifically:

Linear models

The long-term forecasts of the linear and non-linear specifications are given by:

$$E_i[x_{i+s}] = \omega \frac{1-(\alpha+\beta)^{s-1}}{1-(\alpha+\beta)} + (\alpha+\beta)^{s-1}\psi_{i+1} \qquad (4.22)$$

When OTC transactions are taken into account (ACD-OTC) the following formula is used instead.

[114] The Kolmogorov-Smirnov statistic (KS) is a non-parametric test of goodness of fit continuous Probability Density Functions (PDFs) that can be used to compare a sample with a reference probability distribution. KS is:

$$KS = \sup_x |F_n(x) - F(x)|,$$

where $F_n(x)$ is the empirical distribution function, $F(x)$ the theoretical and $\sup x$ is the supremum of the set of distances. The closer the empirical distribution is to the theoretical, the closer KS is to 0. Practically, the test needs large samples to properly reject the null hypothesis.

[115] This methodology is preferred because it relies only on the assumption that the residuals are correctly specified, and not on their distribution. For further information see Dufour and Engle (2000a). However, this methodology is based on the assumption that longer forecasts tend to be closer to the mean. Therefore, the longer the forecasting horizon the closer, asymptotically, to the mean is the forecast. This is a very general approach that summarizes in a very rigid way past information.

$$E_i[x_{i+s}] = \omega \frac{1-((\alpha+\zeta*D_{i-1})+\beta)^{s-1}}{1-((\alpha+\zeta*D_{i-1})+\beta)} + ((\alpha+\zeta*D_{i-1})+\beta)^{s-1}\psi_{\iota+1} \tag{4.23}$$

where $\widehat{\omega}, \hat{\alpha}, \hat{\beta}$ and $\hat{\zeta}$ are the estimated parameters of the respective ACD model.

Non-linear models

Ten-step forecasts for non-linear models are given by:

$$E_i[x_{i+s}] = \xi_1.\xi_2.\dots.\xi_{s-1}\left[exp\left\{\omega\frac{1-\beta^{s-1}}{1-\beta} + \beta^{s-1}\ln{(\psi_{\iota+1})}\right\}\right] \tag{4.24}$$

where $\widehat{\omega}, \hat{\alpha}$ and $\hat{\beta}$ are the estimated parameters from the ACD models and $s = 10$ is the forecasting horizon, and

$$\xi_m = E_{i+s-m-1}[exp\{\beta^{m-1}g(\varepsilon_{i+s-m})\}] \tag{4.25}$$

are parameters that are used to correct biased forecasts (Dufour and Engle, 2000a), and can be estimated using the following sample moment;

$$\widehat{\xi_m} = 1/R\sum_{i=1}^{R} exp\{\beta^{m-1}g(\varepsilon_{i+s-m})\} \tag{4.26}$$

where m=1, … , s-1 and $g(\varepsilon_i)$ is a non-linear function of past standardized durations Appendix 4.A Table 4.2.

Performance Measures

The in-sample fitted values and out-of-sample forecasts, generated by each specification, are tested using the following loss function:

$$UNL = {}^{MSE}/_{\bar{\hat{x}}} \tag{4.27}$$

$$MSE \equiv {}^{1}/_{(N_T+1)}\sum_{i=R}^{T-s}\left(x_{i+s}-\hat{x}_{i+s,i}\right)^2 \tag{4.28}$$

where *UNL* is the unitized loss, which measures the average squared error for forecasted durations, MSE is the mean squared error, x is the realized duration, x̂ is the forecasted duration, and N_T is the number of observations. Unitized Loss (*UNL*) could be a measure of expected loss, or the expected risk when dealing with pricing models, for the average duration.

Moreover, the Correlation Coefficient (*CORR*) between predictions and corresponding realized values is also calculated as a measure of goodness of forecasting accuracy:

$$CORR = {}^{1}\!/_{(N_T - s - 1)} \sum_{i=R}^{T-s} (x_{i+s} - \bar{x}) \left(\hat{x}_{i+s,i} - \bar{\hat{x}}\right). \qquad (4.29)$$

4.3 Empirical Results

The estimation and forecasting results are presented in Appendices 4.B and 4.C. In Appendix 4.B there are four broad tables, each of them with three panels. Tables 4.3, 4.4, 4.5 and 4.6 summarize the estimation results for ECX I, ECX II, NP I and NP II, respectively. Panel A of each table presents the estimation results, the associated statistics and the hypothesis tests for three existing models: ACD, Log-ACD and EXACD. Panel B of each table presents similar information for the ACD-OTC family of models. Panel C of each table presents similar information for the ST-BCACD family of models. E, W and G stand for Exponential, Weibull and Generalized Gamma distributions. In addition, the first section of each table presents the estimation results, where the values in parentheses are the associated *t-statistics*. The next section presents the Log-Likelihood function value, *L*, and the Bayesian Information Criterion (*BIC*). The next section presents the hypothesis testing for the additional parameters in W and G models, where the values in parentheses are the associated *p-values*. The last section presents the Kolmogorov-Smirnov statistic, *KS-stat*, and the associated *p-values*.

Furthermore, Table 4.7 in Appendix 4.C presents the ranking of the models according to their *Q-statistics*. The first column presents the models and the associated distributions. The next three columns report the ranking, the *Q-statistic* and the *p-value* in parenthesis for ECX I. The next three sections present similar results for ECX II, NP I and NP II, respectively. Table 4.8 in Appendix 4C reports the ranking of the models according to their in-sample one-step "forecasts". The first column presents the models along with the associated distributions. The first section reports the rank and the actual value of *UNL* in both markets and phases. The next section reports similar results according to *CORR* loss function. The next two columns present the average ranking of each model in each phase. The next two columns present the average ranking of each model in each market, while the last one reports the average ranking, Total, of each model. Table 4.9 presents similar results according to the out-of-sample one-step forecasts of the models, while Table 4.10 focuses on the out-of-sample ten-step forecasts.

Finally figures 4.1, 4.2, 4.3 and 4.4 present the Q-Q plots of all models in ECX I, ECX II, NP I and NP II, respectively. Panel A of each figure presents the Q-Q plots of existing models. The first column refers to the basic ACD model, the second refers to the Log-ACD, while the last to the EX-ACD. Panel B of each figure presents the Q-Q plots of new models proposed in this study. The first two columns refer to the ACD-OTC and BCACD-OTC models, while the last column refers to the STM-ACD, T-ACD and T-ACD-OTC models. Panel C of each figure presents the Q-Q plots of the BCACD, in the first column, the SEST-BCACD, in the second column, and the STV-BCACD, in the third column. E, W and G refer to the Exponential, Weibull and Generalized Gama distributions.

A quick inspection of Tables 4.3, 4.4, 4.5 and 4.6 reveals a consistency of parameter estimates in the Carbon market with more liquid markets. More specifically, the parameter ω is positive in ACD and Log-ACD models and negative in EXACD. Similar results have been reported by Dufour and Engle (2000a). The range is from 0.0201 in the E-ACD in ECX I to 0.2483 in G-Log-ACD in ECX II. For EXACD ω varies from -0.0494 in E-EXACD in NP II to -0.0982 in W-EXACD in NP I. In general, ω coefficient increases in magnitude when more complex distributions are employed. Generalized Gamma (G) is associated with the largest values. In addition, ω is larger in absolute value when the non-linear, log, specification is employed for the conditional mean instead of the linear-ARMA. Further, this coefficient is on average larger in ECX II and in NP I, which are the most liquid data sets in each market. This raises concerns about the liquidity impact on the performance of the models.

The estimated value of parameter α is also consistent with the literature. It is always positive, varying from 0.1133 (E-Log-ACD in NP II) to 0.4401 (G-EXACD in ECX II). It increases across distributions in a way similar to ω, i.e., as flexibility increases by using additional distribution parameters. Moreover, estimates of the coefficient β, as reported in the literature, are always positive and less than 1. β varies from 0.5811 (G-Log-ACD in NP II) to 0.9751 (E-EXACD in ECX I). Contrary to ω and α, it decreases across distributions. A closer inspection of the tables reveals that in the same family of models β has smaller values for larger ω's and α's. In panel A of Table 4.3, β in the

ACD model decreases from 0.8586, when E is employed, to 0.6970, when G is employed. The corresponding values for ω are 0.0201 and 0.0451. In addition, it is always larger in EXACD, probably because of high negative values of ω. Moreover, it is consistently larger in ECX (e.g., 0.8586 (E), 0.7962 (W) and 0.6970 (G) in ECX I and 0.7334 (E), 0.6880 (W) and 0.6533 (G) in NP) and especially in Phase I (e.g., in ECX is 0.8586 (E), 0.7962 (W) and 0.6970 (G) in Phase I and 0.7892 (E), 0.7114 (W) and 0.5853 (G) in Phase II).

With respect to modelling, almost all parameter estimates presented in the four tables are statistically significant. This is somewhat expected for large samples, such as the one we have for ECX, but is also the case for NP, which has a markedly lower number of observations. In particular, significant estimates for α and β in all tables indicate the underlying autoregressive dynamics, and hence the appropriateness of the ARMA/ACD modelling framework for duration. Furthermore, estimates of α and β for ACWD in all four tables add up to less than, but close to, one. This indicates stationarity with high persistence. Thus, intensity shocks have prolonged subsequent effects, in the sense that a shock does not die out quickly, but it persists over time.

Another parameter of particular interest is coefficient ζ.[116] This parameter captures the effect of extreme durations on the conditional mean. Estimates of this parameter are negative and always statistically significant. This implies that after very long or very short durations the expected duration is more conservative, thus shorter. Specifically, the longer or shorter the duration is compared to the mean, the shorter is the forecast. It is always smaller than α in absolute value and works as a correction factor to both α and ω. ζ increases across distributions like α and ω. It varies from -0.0932 (E-EXACD in ECX I) to -0.3925 (G-EXACD in ECX II). In addition, it appears to be larger in Phase II (e.g., -0.0932 (E), -0.1723 (W) and -0.2340 (G) in ECX I and -0.1690 (E), -0.2578 (W) and -0.9325 (G) in ECX II), especially in ECX (e.g., -0.1690 (E), -0.2578 (W) and -0.9325 (G) in ECX II and -0.2007 (E), -0.2467 (W) and -0.2580 (G) in NP II). The larger parameter ζ in Phase II and in ECX is a sign that the Carbon market is more mature after the completion of Phase I, especially in the larger exchange, because it can absorb duration shocks faster. Moreover, the initial hypothesis that ζ=0 is rejected in almost all markets and phases.

[116] This parameter is named delta in the original work of Dufour and Engle (2000a). In this study the name ζ is used mainly for abbreviation reasons.

The estimation results for the OTC family of models are presented in panel B of Tables 4.3, 4.4, 4.5 and 4.6 and are totally in line with the previously discussed models. In particular, and in accordance with the literature, estimates of ω are always positive in linear specifications and negative in the non-linear BCACD-OTC specifications. It always increases in absolute value when more complex distributions are employed. Its estimated value ranges from 0.0266 (E-ACD-OTC in ECX I) to 0.1432 (G-ACD-OTC in NP II) in the linear model and from -2.0810 (G-BCACD-OTC in ECX II) in the non-linear model to -0.1711 (E-BCACD-OTC in ECX I). In addition, estimates of α are always positive and increase along with higher complexity of the assumed distribution and varies from 0.1007 (E-ACD-OTC in ECX I) to 2.2348 (G-BCACD-OTC in ECX II). In line with parameter estimates of the old models, estimates of β are always positive and larger in Phase I and in ECX, and they decrease across distributions. Lower estimates of β correspond to larger absolute estimates of ω and α. The range of β is from 0.5626 (G-ACD-OTC in ECX II) to 0.9653 (E-BCACD-OTC in ECX I).

The main parameter under examination in this family of models is ζ. It captures the effect of OTC transactions on the conditional mean and works as a correction factor of α for OTC transactions. From a statistical point of view, estimates of ζ, along with estimates of α and ω, increase across distributions. Estimates of ζ range from 0.0250 (E-ACD-OTC in NP II) to 0.2897 (G-BCACD-OTC in NP I) and tend to be larger in BCACD-OTC models (e.g., 0.0509 (E), 0.0736 (W) and 0.1050 (G) in ACD-OTC and 0.0595 (E), 0.0992 (W) and 0.2801 (G) in BCACD-OTC in ECX I). Finally, ζ is always statistically significant, and the null hypothesis that ζ equals 0 is always rejected. ζ is always $0 < \zeta < 1$.

These estimates mean that when OTC EUA holders enter the market, the durations of succeeding transactions tend to be longer. A possible interpretation could be that OTC transactions are assumed to be informative trades and, therefore, represent an information inflow to the market. OTC EUA holders are not obliged to register their positions and when they decide to enter the market, they might do so because they have some reason. In this case other market participants might be reluctant to trade following an OTC transaction, because of potential asymmetric information. They might believe that there is a higher presence of informed traders, or unresolved information in the market. This might deter them from trading. Another possible explanation is that, since usually OTC transactions are large, they might “consume” the current liquidity and the market needs time to reach a new equilibrium. The discussion in Section 5.3 indicates that OTC transactions are likely to be related to information, mainly due to their large size. Along the same lines, the findings in Section 6.3 confirm that OTC transactions decrease the expected trading intensity and have a significant price impact. Consequently, the negative correlation between OTC transactions and expected duration is more likely to be related to information.

Another important parameter estimated here is δ, which measures the size of the impact of the innovations on the conditional mean. δ is always positive and usually less than 1, $\delta < 1$. This implies a milder effect of past durations. Its estimates are slightly higher than in the basic BCACD model (e.g., 0.5434 (E), 0.4081 (W) and 0.1332 (G) in BCACD and 0.5915 (E), 0.4501 (W) and 0.1816 (G) in BCACD-OTC in ECX I). Considering that the models are implemented with the parameter ζ, which is positive, the combined effect is a larger impact of past realized durations on the conditional mean. Consequently, the associated β's are smaller. In addition, similar to β, estimates of δ decrease across distributions and have a range from 0.1190 (G-BCACD-OTC in ECX II) to 0.5915 (E-BCACD-OTC in ECX I). Also, the initial hypothesis that $\delta = 1$ is always rejected.

The next family of models estimated is the ST-BCACD models (please refer to Panel C of tables 4.3, 4.4, 4.5 and 4.6). These models allow the parameter δ, which measures the size of the impact of the past realized durations, to vary across different regimes of the threshold variable (i.e., trading intensity). The magnitude of the coefficient estimates is in line with that reported in the literature. More specifically, estimates of the parameter ω are always negative with higher values across distributions and are always larger in ECX and especially in Phase II (-0.1290 (E-SEST-BCACD in NP I) to -2.293493 (G-STV-BCACD in ECX I)).[117] The same trends are observed in estimates of coefficient α, which are always positive and range from (0.1905 (E-STV-BCACD in ECX I) to 2.5167 (G-STV-BCACD in ECX I)). In contrast, estimates of the coefficient β are always larger in ECX (e.g., 0.9740 (E), 0.9466 (W) and 0.8837 (G) in ECX I and 0.9080 (E), 0.8718 (W) and 0.8258 (G) in NP I), especially in Phase I (e.g., 0.9740 (E), 0.9466 (W) and 0.8837 (G) in ECX I and 0.8943 (E), 0.8696 (W) and 0.8482 (G) in ECX II), and decrease across distributions (0.7676 (G-BCACD in NP II) to 0.9877 (E-SEST-BCACD in ECX I)).[118] Furthermore, estimates of the coefficient δ appear to be

[117] Especially in ECX, whenever the Generalized-Gamma distribution is employed, ω>1. Following Dufour and Engle (2000a), that seems to be consistent.
[118] Another comment that should be made here is the large values of t-statistics, reported in parentheses, for some coefficients, especially β. This is consistent with the literature and it is a result of the deterministic character of the

always less than 1 (e.g., 0.5434 (E), 0.4081 (W) and 0.1332 (G) in ECX I). Hypothesis tests always reject that $\delta = 1$. Its value also decreases across distributions, in a way similar to that of β. In addition, δ is always the weighted average of $\delta 1$ and $\delta 2$, where these two coefficients define a range of possible values depending on the stage that the last observation of the threshold variable is in (e.g., in ECX I δ is 0.5434 in BCACD, while $\delta 1$ is 0.4707 and $\delta 2$ is 1.2394, when E is employed).

When past durations are used as threshold variable and the threshold variable is on a high range, indicating long durations, $sk > \bar{x}$, the size of the effect of past duration of the two regimes is $\delta_1 < \delta_2$. This means that when the previous duration is long, $\delta' \rightarrow \delta_2$ and the longer the duration, the higher the size of its impact on the conditional mean. The economic interpretation is that this model allows for durations longer than the mean to have a larger impact on expected duration, thus increasing its value. This is more profound in Phase I, when, due to the early stage of the market, longer durations were observed. Both parameters are highly statistically significant. This shows that the ARMA specification is not sufficient in capturing a non-linear clustering effect. More specifically, the threshold value is everywhere larger than the mean and it is always larger in NP and in Phase II, sometimes up to four times the mean. Therefore, it seems that as market gains complexity, liquidity increases proportionally. Consequently, inactive stages become rarer and they are captured by a separate regime, where the associated durations are relatively large. In addition, the difference between ECX and NP is emphasized by the large difference in the smoothness parameter, almost eight times as large in Phase II. Although still quite high in ECX I, the transition between regimes seems to be smoother in ECX. This could be considered as an indication of a more sophisticated trading environment, where information dissemination is faster and more efficient, or simply as a natural outcome of the increased liquidity. Indeed, in NP where the relative intensity of trading is considerably lower, distinct regimes are observed, showing potentially sharp adjustments to new information.

Similarly, when trading intensity is used as a threshold variable, $\delta 1 > \delta 2$ and $sk < \bar{x}$. The values of $\delta 1$ and $\delta 2$ vary across a narrower range, compared with simple durations (e.g., in ECX I $\delta 1$ and $\delta 2$ are 0.4707 and 1.2394 (E), 0.3928 and 0.8003 (W) and 0.1286 and 0.1463 (G) in SEST-BCACD, while $\delta 1$ and $\delta 2$ are 1.1861 and 0.4935 (E),

model and a combination of the method of estimation and the number of observations. Another factor that contributes to that problem is the immature state of the market especially in the beginning. This introduces heteroskedasticity of known form. Therefore, all models are estimated using heteroskedasticity robust errors.

0.6833 and 0.4082 (W) and 1.466 and 0.1031 (G) in STV-BCACD). This means that when the market is in an inactive stage, i.e., the last transaction is of a small size, or the associated duration is large, the size of the effect of past durations is expected to be large ($\delta' \rightarrow \delta_1$), and therefore the expected duration is longer. In addition, the inclusion of trade size confirms the previous findings that, in Phase II, ST-BCACD captures the effect of very inactive stages of the market. Indeed, the threshold values in Phase II are always smaller (e.g., in ECX *s* is 0.7783 (E), 0.8838 (W) and 0.9065 (G) in Phase I and 0.4051 (E), 0.4517 (W) and 0.5377 (G) in Phase II), indicating lower volumes and/or longer durations, without exhibiting systematic patterns between the two exchanges. In addition, in both cases, when either duration or trading intensity is used as a threshold variable, the smoothness parameter decreases across distributions, since the increased flexibility of more complex specifications seems to explain partially the DGP of durations (e.g., 1.0625 (E), 1.0150 (W) and 1.0045 (G) in SEST-BCACD and 0.7783 (E), 0.8838 (W) and 0.9065 (G) in STV-BCACD in ECX I).

Furthermore, the analysis of the switching-regime models provides a further insight into the trading process of the Carbon market. Long durations are followed by even longer durations, especially when the associated trading size is low. In contrast, when trading intensity is high, the momentum of the market is expected to continue with fast/large-sized trading. Assuming that investors observe order flow and update themselves continuously, they tend to transact faster and in larger volumes when the market is more active. According to Easley and O’Hara’s (1992) propositions, this is the case when informed trading is more active and supports the initial description of the market as a highly speculative environment. In addition, considering that OTC transactions increase the expected duration, they seem to slow down, in terms of longer durations, and thus lower trading frequency, i.e., the trading momentum. Increased durations might be due either to traders over-reaction to the informational content of OTC trades, or, simply to the fact that these transactions are, on average, very large for the current order book depth. Both the informational and the liquidity content of trades, especially of OTC trades, and their associated durations, as well as their price impact, are further examined in the following two chapters.

Another issue to be discussed concerns the generality and complexity required by the underlying distribution in order to capture higher moment features of the data. First, γ and λ are statistically significant in all estimations. Moreover, Wald hypothesis tests for these parameters are reported at the bottom of each table. The initial hypotheses that

$\gamma = 1, \lambda = 1$ and $\gamma = \lambda = 1$ can all be easily rejected. This indicates that E is insufficient in describing the DGP of duration, while the Generalized Gamma is more appropriate, since λ is statistically different from one. In addition, further analysis of the fitting and forecasting ability of the models confirms this preliminary finding. In particular, estimates of λ range from 1.1632 (G-ACD NP II) to 13.8026 (G-Log-ACD ECX I) and are, therefore, much higher than 1, although markedly more so for ECX than for NP, indicating the marked effect of large observations. Furthermore, in the Weibull models $0 < \gamma < 1$ always. This indicates a monotonically decreasing hazard function, which means that the probability of a transaction to occur decreases with time. Furthermore, in G, $0 < \gamma < 1$ and $\lambda > 1$ always. In ECX $\gamma^*\lambda>1$ and this implies a non-monotonic, inverse U-shaped hazard function. In contrast, in Nord Pool $\gamma^*\lambda < 1$ always, which, according to Lunde (2000), indicates a monotonically decreasing hazard function. This is strong evidence of different structures across market and differences in trading patterns.

More importantly, the maximum Log-Likelihood function value (L) reported in the middle panel of the tables is significantly greater for models estimated with the Generalized Gamma (G) distribution than with the Weibull (W) distribution than with the Exponential (E), in this order.[119] For example, in Table 2 the optimal Log-likelihood function values, L, are -21576.69 for W and -20828.69 for a G. These values give a Likelihood Ratio test statistic of 2651.74, which is far larger than the 5 percent critical value of 3.84. As far as this statistic is concerned, this indicates that the ACD model specifications, which assume that the Data Generation Process of duration follows a Generalized Gamma distribution is significantly better, both statistically and on the basis of likelihood, than the one with the Weibull distribution. This result is consistent throughout the tables for all estimated, both linear and non linear, models. Thus, Generalized Gamma distribution provides a statistically better estimation in both markets and both phases and across linear and non-linear models.[120]

Focusing on the second section of tables 4.3, 4.4, 4.5 and 4.6, as well as on Table 4.7, the *BIC* and *Q-stat(15)* confirm these results. *BIC* is consistently smaller when more

[119] The nesting property of the Generalised Gamma distribution to the Weibul and the Exponential, and that of the Weiibul to the Exponential facilitates testing using the Likelihood Ratio, which is distributed as chi-squared with a degree of freedom equal to the difference in the number test statistics of parameters between pairs of models.

[120] Strictly speaking, we cannot use the Likelihood Ratio to test whether the non-linear BCACWD-OTC fits better than the ACWD-OTC as the former does not directly nest the latter. However, comparison is conducted later on the basis of in-sample forecast performance.

complex distributions are employed. Similarly, the autocorrelation of the squared standardized durations is consistently lower when higher flexibility is employed (e.g., 8.16 (E), 6.85 (W) and 5.51 (G) in the basic ACD model in ECX I). Consequently, Generalized Gamma distribution appears to be more appropriate in describing the DGP of duration, compared to its nested alternatives. The extra flexibility adds significantly to the fitting of the models. This can be also seen in Table 4.8, where more complex distributions achieve a higher in-sample one-step "forecasting" accuracy (e.g., *UNL* (lower is better) is 6.6977 (E), 6.7994 (W) and 6.5263 (G), while *CORR* (higher is better) is 0.3904 (E), 0.4181 (W) and 0.4374 (G) in the basic ACD model in ECX I). However, as Bauwens et al. (2004) argue that simple models, mainly referring to simple distributions, are inadequate to capture duration series peculiarities.[121] However, the main determinant of the DGP of duration still remains the conditional mean specification, although more complex distributions consistently improve fitting. Therefore, market peculiarities, such as OTC transactions or liquidity patterns, appear to be more important stylized facts that need to be modelled first.

Focusing on the in-sample performance of the models, Table 4.7 presents the ranking of the models according to their ability to reduce the remaining autocorrelation in the standardized duration series. The higher ranking of linear specifications, according to lower *Q-stat(15)* (e.g., the ranking of ACD is 4 (E), 3 (W) and 1 (G), while the ranking of ACD-OTC is 9 (E), 7 (W) and 6 (G) in ECX I), supports their superiority against their non-linear counterparts. Although all models significantly reduce autocorrelation, non-linear models fail to surpass their linear counterparts (e.g., STV-BCACD ranking is 20 (E), 15 (W) and 10 (G) and BCACD-OTC ranking is 21 (E), 18 (W) and 12 (G) in ECX I) and in some cases they still leave some unexplained autocorrelation (e.g., Log-ACD ranking is 26 (E), 27 (W) and 25 (G) and BCACD ranking is 24 (E), 19 (W) and 11 (G) in ECX I). In addition, OTC models significantly improve fitting in NP (e.g., ACD-OTC ranking is 2 (E), 5 (W) and 8 (G) in Phase I). In contrast, the regime-switching framework appears to be more relevant in ECX and in Phase I in particular. The ranking of G-SEST-BCACD is 13, while the relative value for STV-BCACD is 10.

Along the same lines, Table 4.8 presents the in-sample short-term "forecasts" for all estimated models. The main findings confirm the superiority of linear models,

[121] The analysis presented in the next chapter indicates that a mixture of distributions improves the, already good, fitting of Generalized Gamma. However, the main determinant of describing the DGP of duration is the conditional mean equation specification.

especially when G is assumed to be the associated density function.[122] The linear specifications, such as the basic ACD (e.g., in ECX I the ranking is 3 (E), 5 (W) and 2 (G) according to *UNL* and 9 (E), 7 (W) and 2 (G) according to *CORR*) and the ACD-OTC (e.g., in ECX I the ranking is 7 (E), 9 (W) and 10 (G) according to *UNL* and 8 (E), 6 (W) and 4 (G) according to *CORR*) models, achieve better fitting and get better ranking compared to their non-linear counterparts, such as the Log-ACD (e.g., in ECX I the ranking is 21 (E), 12 (W) and 8 (G) according to *UNL* and 24 (E), 23 (W) and 22 (G) according to *CORR*) and the BCACD (e.g., in ECX I the ranking is 22 (E), 20 (W) and 14 (G) according to *UNL* and 18 (E), 16 (W) and 14 (G) according to *CORR*) models. The density function specification almost everywhere improves fitting of the same conditional mean specification (i.e., G provides consistently better results than W, which in turn provides higher flexibility, and thus higher accuracy, than E), but the latter is the major factor in describing the DGP of duration. As can be seen in Table 4.8, almost everywhere the conditional mean specifications improve or worsen the performance of a model noticeably, while more complex distributions only slightly increase accuracy. This is profound when OTC transactions are taken into account, where the ranking of the models improves significantly. These transactions appear to be more relevant to NP and especially in Phase I, where ACD-OTC and T-ACD-OTC are among the best performing models.

Another finding is that transaction size improves the fitting of ST-BCACD models compared to durations, thus indicating that size does matter. The STV-BCACD model, especially when G is assumed to be the distribution of durations, consistently provides improved forecasts (e.g., according to *UNL* the ranking of G-STV-BCACD is 13 (ECX I), 7 (ECX II), 20 (NP I) and 12 NP II, while according to *CORR* is 10 (ECX I), 12 (ECX II), 18 (NP I) and 11 (NP II)) compared to the SEST-BCACD. A further analysis of the role of trading volume in duration modelling and its connection to information will be undertaken in the following two chapters.

These findings are confirmed by the KS-statistics, reported in Tables 4.3, 4.4, 4.5 and 4.6, and the Q-Q plots, presented in figures 4.1, 4.2, 4.3 and 4.4. Generalized Gamma models always improve fitting, but the main improvements come from the conditional mean specification. The nested distributions, especially E, do not fit particularly well,

[122] The STM-ACD and the T-ACD models are still linear parameterizations, at least in every regime. These models are presented in the following chapter. Their fitting and forecasting accuracy is compared to the models discussed in this chapter and therefore they are all included in the same tables in appendices 4.B and 4.C.

since there are significantly observable discrepancies in the long right tails of the empirical and theoretical distributions, which is captured by the large deviations in the upper end of the straight diagonal line of the Q-Q plots. The first line of panels A, B and C of figures 4.1, 4.2, 4.3 and 4.4 show that the empirical distribution significantly deviates from the theoretical. This means that these distributions cannot explain the very long or very short durations, either because they overestimate or underestimate their probability of being observed. Therefore, the need for the higher flexibility of G is visually obvious. In addition, the new models proposed outperform the existing models, which, especially Log-ACD, EX-ACD and BCACD, do not fit particularly well.

In contrast to the in-sample forecast results of Table 4.8, the short-term out-of-sample forecasts of Table 4.9 indicate that, linear models, fail to capture the short-term dynamics needed to forecast the next duration even though they have a reasonably good in-sample fit. For example, in ECX I the ranking of the basic ACD model is 20 (E), 21 (W) and 19 (G) according to *UNL* and 19 (E), 17 (W) and 15 (G) according to *CORR*, while the ranking of ACD-OTC is 23 (E), 24 (W) and 22 (G) according to *UNL* and 18 (E), 16 (W) and 14 (G) according to *CORR*. The regime-switching framework seems to provide pretty accurate forecasts, especially in Phase I in ECX (e.g., G-STV-BCACD ranking is 7 according to *UNL* and 4 according to *CORR*). In contrast, OTC models perform better in Phase II (e.g., the average ranking in Phase II is 4 for the G-BCACD-OTC) and in NP in particular (e.g., the average ranking of the G-BCACD-OTC is 7 in NP). In addition, the volume enhanced durations, using trading intensity, significantly improve the performance of ST-BCACD, which are always among the best performing models in all markets and phases. The total ranking of G-STV-BCACD is 4, and it is the best performing model after the ones presented in Chapter 5. These results provide the motivation and the foundation for the analysis of the following chapters, where the importance of trading intensity is analyzed further. Finally, the Generalized-Gamma consistently provides the most accurate forecasting and seems to be the most appropriate density function among the three considered.

Table 4.10 presents long-term out-of-sample forecasts, and these raise again the importance of linear specifications. The results in the table show that, although G is still sufficient, the non-linear specifications fail to provide accurate forecasts. In some cases, such as in EX-ACD (total ranking is 18 (E), 20 (W) and 14 (G)) and BCACD (total ranking is (24 (E), 27 (W) and 25 (G)), the *UNL* provides pretty inaccurate results. In contrast, linear models, especially the OTC specifications, improve long-term forecasting. OTC models are always better (that mainly refers to the T-ACD-OTC, discussed in Chapter 6), but, compared to the regime-switching models, the latter rank better according to *CORR* (e.g., the ranking of G-STV-BCACD is 21 (ECX I), 17 (ECX II), 21 (NP I) and 20 (NP II) according to *UNL*, and 7 (ECX I), 11 (ECX II), 5 (NP I) and 25 (NPII) according to *CORR*). This means that although the regime models are inaccurate in magnitude (i.e., *UNL* is large), they can forecast direction much better (i.e., *CORR* is high).

Finally, the superiority of OTC models concerning long-term forecasts provides a further insight into the Carbon market. The regime-switching framework performs better in the short term, especially in ECX, while OTC transactions play a more crucial role in the long term, especially in NP. This could be interpreted as a strong evidence of structural differences between these two markets. Investors, especially in ECX and in Phase I, seem to update themselves continuously by observing order flow variations, and they adjust their trading accordingly. This is why the regime framework has a slightly better performance in this market, especially when past trading history, such as past trading intensity, is taken into account. The net effect is that order flow has a momentum impact on trading patterns. Therefore, $\hat{\delta}_1$ and $\hat{\delta}_2$ (tables 4.3, 4.4, 4.5 and 4.6) indicate that when trading intensity is high, the expected duration is shorter. Easley and O'Hara (1992) argue that high trading activity could be interpreted as informative. This idea will be further developed in Chapter 5. In contrast, when an OTC EUA holder enters the market, the trading process seems to slow down, either due to asymmetric information across market participants or due to shortage of liquidity. Whichever the reason, it seems to have a long-term effect, and thus a better long-term forecasting, by increasing the expected duration. Consequently, these preliminary results, given the speculative character of the market, could practically mean that traders in the Carbon market are quite sensitive to new information and thus order flow creates a momentum, while OTC transactions are considered to carry information that is not immediately resolved and, thus, these transactions have a prominent role in the long term.

4.4 Summary

When analytical interest focuses on the intraday formation of various microstructure phenomena, special attention is required concerning time deformation. Economic events are randomly spread throughout the trading day and the intermediate time, defined as duration, is irregularly spaced. Therefore, several econometric models might not hold unconditionally. In a seminal study, Engle and Russell (1998) propose the ACD model for modelling duration clustering and over-dispersion. The ACD framework models time exogenously as a dependent stochastic process. Modelling requires a conditional mean specification conditional on past durations and a density function with positive support. The exogeneity assumption has recently been debated and it will be further discussed in Chapter 6.

The ACD class of models bears a very strong resemblance to GARCH and therefore has been developed rapidly and become very popular. The main objections against the basic ACD model of Engle and Russell (1998) focus on the appropriateness of the conditional mean specification and the distribution assumed for the standardized durations. Asymmetric effects and non-linear modelling of past durations, in particular, have been an area where the literature is rich. In addition, several distributions have been proposed. However, it seems that there is no relevant study in the Carbon market, which exhibits various structural differences compared to other well established, and far more liquid, markets.

In the present study, the analysis focuses on duration in the Carbon market, paying special attention to non-linear effects and to the impact of OTC transactions. In more detail, the fitting and forecasting of several existing models, such as the original ACD, the Log-ACD, the BCACD and the EXACD, are examined, employing the Exponential, the Weibull and the Generalized-Gamma distributions. Model fit is tested by the ability to reduce the autocorrelation of squared standardized durations, the appropriateness of the associated density function and short-term in-sample forecast accuracy. The forecasting performance is determined by their short- and long-term out-of-sample forecasting ability.

In addition, two empirical extensions are proposed. The first introduces a highly non-linear term in the existing Dufour and Engle's (2000a) BCACD model. The size of the impact of past durations is allowed to vary with changes in economically relevant threshold variables. The threshold variables chosen are past durations (i.e., self-exciting process) and trading intensity (i.e., volume enhanced durations). Terasvirta's smooth transition regime-switching framework is employed allowing for gradual adjustment to continuously changing market conditions. The second extension incorporates the duration impact of OTC transactions. OTC EUA holders are allowed to enter the market, and this is expected to have an impact on the DGP of duration. Therefore, a new term is introduced that allows the effect of past durations to be revised when the last transaction is an OTC. The ARMA and the BCACD specifications are employed to account for both linear and non-linear effects.

One of the main findings refers to the relative performance of the various specifications. Although more complex distributions are required to better describe the data’s higher moments, fitting and forecasting are mainly influenced by the particular conditional mean specification assumed. The Generalized Gamma performs consistently better than the other two nested distributions, while Exponential provides the worst forecasts and fitting, but the modelling of particular market peculiarities, such as the OTC transactions, is of particular importance, and improves performance significantly. This confirms the initial propositions by Bauwens et al. (2004) that simple models, especially the ones with a single distribution, are inadequate to capture duration peculiarities. The modelling of OTC transactions and of asymmetric effects appears to improve performance. Linear models fit the data better and provide better long-term forecasts. However, the short-term dynamics are better described by non-linear models, because they provide more accurate one-step forecasting.

The switching regime analysis indicates that there are blocks of high or low trading activity in the market and that order flow creates a momentum in trading. Given Easley and O’Hara’s (1992) propositions concerning the presence of informed trading, these high trading activity blocks could be interpreted as information episodes. These episodes appear to be particularly relevant in Phase I, especially in ECX. They also provide better short-term forecasts. In contrast, OTC transactions increase the expected duration, either due to information reasons or because they deplete liquidity. These models are particularly relevant in NP and especially in Phase II and provide better long-term forecasts.

The relation between trading intensity regimes and OTC transactions provides an insight into the Carbon market. The analysis of the regime-switching framework indicates that after a large, or fast (i.e., short duration), transaction trading frequency increases, while the opposite occurs when trading intensity is low. Therefore, given the speculative environment in the market, investors appear to extract information from previous trades and this has an impact on their trading frequency. This is consistent with Engle and Russell (1998) and Dufour and Engle (2000a, 2000b), who support that duration is exogenously determined. Yet, OTC transactions, since they have an increasing effect on duration, seem to have an opposing effect to market momentum. They either deter other market participants from trading due to adverse-selection or they consume the current order book depth, so the market needs some time to reach a new equilibrium. Whatever the reason, they provide better long-term forecast, while trading intensity regimes provide better short-term accuracy. In practice, this means that investors seem to continuously update themselves by observing order flow variations, which appear to have a short-term impact on trading. In contrast, OTC transactions seem to carry information that cannot be resolved immediately and therefore has a long-lasting effect. Both concepts are relevant to the analysis in the next chapter, where they are further developed.

Chapter 5

Duration Modelling, Trading Intensity and Trade Type

A Smooth-Transition-ACD Model and The Informational Content of Trades

5. Duration Modelling, Trading Intensity and Trade Type

5.1 Introduction

Continuing the analysis of duration modelling in the Carbon market, the analytical focus of this second empirical chapter lies in the theoretical development of ACD models, concerning the distributional assumptions and their connection to the informational content of trades. Following the remarks by Bauwens et al. (2004), who argue that a single distribution cannot summarize the dynamics of durations, several studies develop more sophisticated density functions which try to summarize the presence of different market participants. Consequently, a branch of literature has been developed that connects ACD modelling to the presence of information or informed trading in the market (see, inter alia, Wong et al., 2009; Tay et al., 2009; De Luca and Galo, 2009).

More specifically, the propositions by Bauwens et al. (2004) have given rise to a new statistical approach, which allows the distribution of durations to be described as a mixture of distributions. This offers researchers great flexibility in modelling duration series, allowing them to take into account various issues and stylized facts present in each market. The basic underlying assumption in these models (Hujer and Vuletic, 2007) is that in every market there are different groups of traders that react differently to the arrival of new information.[123] Therefore, their trading behaviour cannot be captured by one single distribution without losing relevant information. Consequently, the collective distribution of duration is assumed to be a mixture of distributions, each of which describes more accurately one group of market participants. Thus, in order to distinguish among the different groups, the researcher needs to define a benchmark variable, the values of which indicate different stages. These stages determine different regimes, which can then be related to one or more groups. The transition mechanism should also be defined.

According to the specification employed, the "mixture distribution" models can be divided into two broad categories. First, De Luca and Zuccolotto (2006) propose a

[123] An important characteristic of such traders' actions is the arrival time of their transactions. The microstructure literature (see, inter alia, Bauwens et al., 2004; Gerhard and Hautsch, 2000; Hujer and Vuletic, 2006) connects the inter-trade arrival time with the type of traders. Since different types of traders are present in the market simultaneously, each transaction could come from a different Data Generation Process (DGP). Consequently, one single distribution should not be sufficient to capture all different types of trades/traders.

switching regime Pareto ACD model using Teräsvirta's logistic smooth transition function, where the parameter of the Pareto distribution is allowed to vary across different regimes of an economically relevant variable. This variable is observable, and when its values exceed or lie between specific benchmarks, namely threshold values, different regimes of that variable are determined. These regimes can be further connected to specific market characteristics, to distinguish among different trading groups. This idea is implicitly assumed in the models of Zhang et al. (2001) and by Meitz and Teräsvirta (2004, 2006). Zhang et al. (2001) use a threshold variable to distinguish between various regimes of the conditional mean specification (ACD), where a different distribution is attributed to each regime; Meitz and Teräsvirta (2004, 2006) allow for a smooth transition. Therefore, these studies recognize different groups of transactions, and thus different regimes, but they do not connect this dichotomy to information.

In contrast, in another class of models the threshold variable used is latent, and therefore unobservable. Thus, although different regimes of this variable cannot be observed directly, they can indirectly indicate the presence, or absence of the factors that cannot be observed but are assumed to have an impact on durations. Following this branch of the literature some authors have proposed a discrete mixture of distributions (e.g., Liu et al, 2004; De Luca and Gallo, 2004) of durations, while some others have proposed a continuous mixture of distributions (e.g., Bauwens and Veredas, 2004; Ghysels et al., 2004). Furthermore, Hujer et al. (2002) propose the Markov-Switching ACD model, which, as they show, has higher forecasting accuracy. These models describe the idea that, although there is no direct way of distinguishing what the motivation of a trade is, the probability of belonging to one regime or another can be defined more accurately.[124] However, these models come at a high computational cost.

As a compromise, Hujer and Vuletic (2007) propose the so-called Static Mixture ACD model, which shares a strong resemblance to Deluca and Gallo's (2004) discrete mixture model, but provides better forecasts. Hujer and Vuletic (2007) allow the DGP of duration to be a discrete mixture of predefined distributions with regimes determined

[124] In Markov-Switching ACD models, the researcher first needs to specify the conditional mean and then the number of the available regimes of the latent variable and the density function associated to each regime. Since the threshold variable is unobservable, different regimes cannot be observed either. Therefore, the probability of moving from one regime to another is of utmost importance, and, depending on the pre-specified number of regimes, a matrix is formulated, consisting of all possible pairs of distributions. The cumulative distribution, i.e., the uniform distribution describing the Data Generation Process of duration, is then the average of all distributions, weighted by the ex ante probability of duration coming from each regime. For more information see Liu et al. (2004) Hujer and Vuletic (2005).

by a latent variable. Each regime is associated with a different mean equation specification for the expected duration in a way similar to the T-ACD proposed by Zhang et al. (2001). In their analysis, they draw inferences concerning informed trading based on the shape of the hazard functions of each regime's distribution.

The microstructure literature recognizes two main types of traders according to their access to price-relevant information (see, inter alia, Kyle, 1985; Glosten and Milgrom, 1985; Easley and O'Hara, 1987, 1992; Easley et al., 1997, 2002). The first group, often called informed traders, possesses private information before it becomes public knowledge. They exploit their timing advantage in order to benefit from subsequent price changes. The second group, referred to as uninformed traders, does not have any prior knowledge of public information, and is assumed to trade because of liquidity reasons. More recent studies dissect the group of uninformed traders further according to their ability to extract information from trading variations and their flexibility in managing transaction size or frequency (e.g., Admati and Pfleiderer, 1988; Foster and Viswanathan, 1990; Spiegel and Subrahmayam, 1992; Gerhard and Hautsch, 2007). They argue that there is a subgroup of uninformed traders that, although they do not possess any price-relevant information, can choose when and how to transact. They are called discretionary liquidity traders and they are assumed to learn from information signals revealed by trades or order flow. Applied to the Carbon market, these three types can distinguish companies which need allowances for compliance, as non-discretionary uninformed liquidity traders, who, consequently, are unrelated to information, market participants, who possess private information, as informed traders and investment funds that try to extract it from order flow, as discretionary-liquidity uninformed traders.[125]

Therefore, the trading pattern, and thus the arrival time of transactions of each trading group should be expected to be largely distinct. More specifically, Hujer and Vuletic (2007) base their inferences on the fact that transactions of uninformed traders are motivated by liquidity rather than by information. Their transaction rate should be invariant over time, since they are not affected by randomly arrived information shocks. Statistically, a flat hazard function could summarize their trading activity. This means

[125] As it has been mentioned earlier, the Carbon market is a cap and trade market, where quantities and prices are indirectly politically influenced. In addition, especially in early stages, OTC market was dominant and transactions were really informative. Although the relative proportion of these trades has declined over the years, OTC allowance holders can still enter the market and their trades are still seen as carrying informational content. Viswanathan (2010) raises the importance of regulation, with regards to information innovation and information dissemination, emphasizing on the role of the OTC market. Therefore, informed traders are expected to play a dominant and profound role in the European Carbon market.

that the Exponential distribution, which was employed by Engle and Russell (1998), should be appropriate for this group. In contrast, information-related trading should follow information innovations, which are usually assumed to arrive randomly in the market. In addition, Easley and O'Hara (1992) support the view that informed traders appear in the market only when there is price-relevant information. Consequently, the transaction rate of informed trading is expected to vary with information episodes., therefore a non-monotonic hazard function, a characteristic of more complex distributions, would be more appropriate. Moreover, Gerhard and Hautsch (2007) extend the idea by adopting the view that portfolio managers and fundamental trading should exhibit an upward slope hazard function due to the fact that they are willing to change the composition of their portfolio only when they aggregate sufficient information.[126] In contrast, investors who acquire or accumulate information, either privately or by extracting it from observing order flow variations, are more likely to trade closer to, or shortly after the arrival of, this new information. Therefore, the probability of short durations is higher and the trading rate of these traders is expected to follow a downward slope.

Furthermore, the mixture of distributions framework provides a better in- and out-of-sample accuracy, but it implicitly assumes a simplified division of market participants that is not affected by market depth or learning speed. However, several studies examine the implications of different transaction sizes and frequencies, as well as of the learning process of market participants.

First, several studies emphasize the importance of trading intensity. The sequential and strategic models of Glosten and Milgrom (1985) and Kyle (1985), respectively, examine the trading pattern variations when informed traders are present in the market and how their appearance is connected to trading intensity. As has been extensively discussed in the Literature Review section, in Chapter 2, Glosten and Milgraom (1985) maintain that informed traders transact immediately upon the arrival of new information and in large quantities in order to fully exploit their information advantage. In contrast, Kyle (1985) argues that they segment their trades, acting strategically, in order to maximize their profits without revealing their information. Apart from their different approach, both

[126] These traders are assumed to be naive and to have an optimal portfolio composition. They are willing to change it only when market conditions impose price changes that are higher than their benchmark. Therefore, their transaction rate, according to information arrival, should have an upward slope because it is more likely for them to transact after gathering enough information. The relative benchmark that indicates enough information is an individual choice and refers to signal strength and not to time.

maintain that the presence of informed traders could be identified, either through higher trading volumes (i.e., large informed transactions) or through higher trading frequency (i.e., many small transactions in a short period of time). Furthermore, the time dimension is underlined by Back and Baruch (2004), who postulate that different traders should be expected to trade at different rates.

However, the time dimension is only implicitly assumed in these two studies. The importance of time is explicitly taken into account in Diamond and Verrecchia (1987). They postulate that longer durations should be associated with bad news, especially in the presence of short selling constraints. In this case, informed traders should sell the asset, but only when they own it. In contrast, Easley and O'Hara (1992) postulate that increased trading activity is associated with higher presence of informed trading. When new information arrives at the market, they possess that piece of knowledge prior to the majority of traders, and they want to exploit it, acting either at once or strategically, before it goes public. This is reflected in increased activity, in the form of shorter durations or of higher volumes. Consequently, both studies emphasize the importance of trading intensity and connect it with information and different types of traders. They provide a qualitative explanation to clustering, which is subsequently modelled explicitly by Engle and Russell (1998) through ACD models.

More recently, various studies emphasize the importance of both trading volume and frequency in the trading process. The literature is settled in many aspects. Dufour and Engle (2000b) connect increased trading activity to higher price impact, confirming Easley and O'Hara (1992). Along the same lines, Ben Sita (2010) suggests that time is an important aspect of intraday trading strategies. Bowe et al. (2007) support this proposition arguing that investors should manage both the time and the size of transaction. In contrast, Grammig et al. (2007) provide evidence of more informative prices after inactive stages of the market, thus confirming Parlour (1998) and Foucault (1999). In addition, several studies, such as those by De Jong et al. (1995), Huang and Stoll (1997), Ahn et al. (2002) and Angelidis and Benos (2009) raise the importance of trading volume for prices and spreads.[127]

All these studies, especially the ones that approach theoretically potential information signals of trading intensity, model intraday trading activity under the assumption that

[127] The relation of transaction size, trading frequency and intraday formation of prices and spreads will be further developed in the next chapter. In the analysis of this chapter these measures of trading intensity are mentioned to emphasize the informational content of trading signals.

market participants observe order flow to extract price-relevant information. In the early literature, inventory-holding models maintain that market makers observe continuously the market to identify informed trading in order to post regret free prices (see, inter alia, Garman, 1976; Amihud and Mendelson, 1980; Stoll, 1978; Ho and Stoll, 1981, 1983; O'Hara and Oldfield, 1986; Hasbrouck, 1988, 1991a, 1991b). They try to formulate expectations concerning the future trading intensity based on the size, the frequency and the direction of previous transactions. Their main concern is market viability and the risk aversion of the dealer. Consequently, the liquidity component of the transactions, which is connected with transitory effects, is taken into account. In contrast, information models see the trading process as a game where players learn from each other by observing past transactions (see, inter alia, Kyle, 1985; Glosten and Milgrom, 1985; Easley and O'Hara, 1987, 1992; Admati and Pfleiderer, 1988; Kyle, 1989; Easley et al., 1997, 2002; Easley et al., 2008). Then, using the acquired knowledge, they formulate expectations of the probability of informed trading (PIN) and take it into account when quoting their prices.

Another important aspect that is discussed extensively in the literature is the definition of learning and the speed of information dissemination.[128] Although market participants observe the same, external, public information, such as the trading process, their prior knowledge (e.g., market share) drives them to interpret the order flow signals in different manners. In addition, their access to private information and the speed of its dissemination varies considerably and, hence, a more dynamic approach should be more relevant. REE literature (see, inter alia, Shiller, 1981, 1984; De Long et al., 1990; Chamley, 2003; Sandroni, 2005) recognizes that learning is a dynamic process, and that, apart from the "actual" truth, there are various perceptions, derived by trading signals that might turn traders, who ignore their private information, towards the "right" or "wrong" direction (see, inter alia, Banerjee, 1992; Bikhchandani et al., 1992; Smith and Sorensen, 2000). According to Vives (2008), "actions speak louder than words" and traders, although they might have their own, potentially correct, piece of information, might "herd" towards an unexpected, potentially wrong signal. However, the way they

[128] The models of Kyle (1985) and Glosten and Milgrom (1985), which set the scene for a series of similar models, belong to the Rational Expectation Equilibrium (REE) models and share a very strong resemblance to Game Theory. They model price formation as equilibrium of expectations formulated by knowledge gained by observing various aspects of order flow variations. Their analytical focus lies on the dynamics of learning and formulating price-relevant expectations. REE literature has been developed in parallel with microstructure literature, and this study by no means aims at developing both. However, several studies might be relevant and therefore they will be briefly discussed. For an extensive discussion please refer to Burrnnermeier (2001) and Vives (2008).

interpret this signal, as well as the speed with which they accumulate and disseminate the information through their trading might vary considerably (inter alia, Vives, 1993; Jun and Vives, 2004).

This branch of literature raises two important points relevant to the present study. First, along with inventory and information models, it recognizes that there are price-relevant information signals that occur before, without necessarily including, price changes. More specifically, market participants in these models observe order flow variations and then try to extract signals that might indicate price changes. There is a time element that is crucial to this approach. Several studies report that different levels of trading intensity are more (less) informative and result in greater (smaller) subsequent price changes (see, inter alia, Easley and O'Hara, 1992; Dufour and Engle, 2000b). Consequently, when traders observe excessive trading volumes or higher trading frequency they get a signal that a significant price change, which could prove beneficial, might follow. They assess these signals and combine them with qualitative information and then they implement their expectations through trading strategies.[129]

The second issue extends the afore-mentioned proposition that each trader's learning speed varies with various factors (see, inter alia, Vives, 1993; Jun and Vives, 2004). Each trader, or group of traders, exhibits a different learning curve, depending on his/her access to private information, risk aversion, trading strategies (e.g., Kyle, 1985; Glosten and Milgrom, 1985) or volume of information. All traders receive the same quantity and quality of public information, but their actions vary in terms of what they learn and how fast they learn from it. [130] Townsend (1978), Frydman (1982), Blume and Easley (1984, 1998), Feldman (1987) and Vives (2008) stress the difference between learning about an equilibrium and learning from an equilibrium. Predictions and expectations can only be formulated in the first case because there is still unresolved information in the market, while in the latter case prices already include all relevant information. Obviously, considering that any information episode has a finite life before

[129] Several studies (Banerjee, 1992; Bikhchandani et al., 1992) report that learning is a dynamic process and is not necessarily Bayesian. Market participants that observe market variations, update themselves continuously and then adapt their strategies accordingly. The social learning literature emphasizes the possibility of market failure due to the fact that fully rational traders may "herd" in the wrong direction because they overestimate/underestimate trading shocks, ignoring any qualitative analysis and their previous knowledge. The decision making depends on the number of unknown parameters, for which expectations need to be formulated, and on the current market dynamics, such as the number of groups of traders and/or relative measures of market activity and depth.

[130] For example, according to Gerhard and Hautsch (2000) there are two types of traders who aggregate information but do not act in the same way. Technical traders react to every signal, while portfolio managers aggregate relevant information up to a level necessary for altering the optimal composition of their portfolios.

information goes public, the time dimension of information plays an important role. First, informed traders have an incentive to act fast or in huge quantities in order to exploit their informational advantage. Their actions reveal their presence to traders that observe order flow. However, not all traders act in the same way or at the same time. They might need either sufficient, aggregated information, in case they manage portfolios, or time to act, due to their different learning speeds. The same happens with non-discretionary liquidity traders, who are the last to adjust to new market conditions and do so mostly when information is known to everyone.

This appears to be particularly relevant to the trading practice in the Carbon market, where information resolution appears to play an important role both in how informative prices are and in liquidity levels, which determine market depth. Viswanathan (2010) argues that the Carbon market needs a regulatory approach that would strictly regulate the market in order to restrict manipulation. Simultaneously, it should allow exchange and market participant innovations to enhance liquidity. A non-regulated, non-transparent market would be liquid, but inaccurate in terms of price. In contrast, a strictly regulated environment would increase price accuracy, but not liquidity. Both would result in a divergence from the initial purpose of the EU ETS. Viswanathan (2010) is based on a seminal study by Kyle and Viswanathan (2008), who argue that organized markets improve the aggregation of diverse information, which, according to Hayek (1945), leads to an increased "market efficiency". However, this does not necessarily mean that it increases "price accuracy" as well. Both studies explain that a highly manipulated market in terms of excessive volume or informed market participants might be highly efficient in incorporating manipulation into price, but assets are not necessarily priced accurately. In terms of Carbon pricing, over- or under-pricing would lead to a highly speculative environment, which would undermine the efforts of emissions reduction.

Both studies provide a solid ground for a further discussion of the link between market participants, trading volume, accurate pricing and regulation. They argue that particular market participants can manipulate the market, and therefore prices, through increased trading volume. This is consistent with the approach by Easley and O'Hara (1992) regarding the time dimension between informational signals, trading volume and subsequent price changes. In addition, they implicitly assume that not all traders act simultaneously and that there is a time difference in accumulating all necessary relevant information. In contrast, the ACD models that employ a mixture of distributions to capture different groups of traders do not account for the quality and speed of learning. They only recognize different types of traders and attempt to capture their trading behaviour by the shape of the hazard functions of the associated durations. In addition, they do not account for non-pricing signals other than the duration itself and they do not allow for gradual information resolution.

This is what this study primarily tries to capture. Price resolution is recognised to be a process that varies over time, not an immediate market adjustment, whereby prices immediately reflect information. In fact, every piece of information needs some time to be fully incorporated into prices. In this period of time, there is a "time dimension" that makes information exploitable.[131] Traders can benefit from extracting this pieces of information during this period of time, since they can trade upon it and be make profits. This way, although they are initially uninformed, they can become informed, compared to other uninformed traders, who do not yet possess that information. Obviously the strength, the direction and the learning process of market participants are of outmost importance. In addition, since prices do not fully incorporate all available information immediately, these traders need to observe "non-price" information signals. This analysis examines whether trading intensity can reveal variations in price-relevant information levels.

This contributes to the literature in various ways. First, the "time dimension" of information is highlighted and explicitly modelled. Prices are recognized to be efficient, in the sense that rationally incorporate information, but imperfect, in the sense that they do not immediately reflect all relevant information. This setting might be too restrictive in a macroeconomic context, but it is rather reasonable on an intraday level, where order submission can be recorded in milliseconds. Second, the ACD framework is related to information in a way that explicitly models the impact of an exogenous, observable variable on trading patterns. Variations in this observable, non-latent, variable can measure how informative a trade is, or has been, based on the hazard function of durations. Consequently, if market participants can forecast future values of this variable, they can have an estimate of how informative a trade is expected to be. This could have numerous practical applications, and could be used as an alternative to the PIN measure proposed in the literature. In addition, this framework can be further

[131] Even in the case of public information, not all trades are executed immediately. Trades are matched according to price and time of order submission. Even a fraction of a second can be exploitable in submitting an appropriate order on an appropriate price. Consequently, high frequency trading strategies that focus on the "time dimension" of information could be profitable.

extended to include other non-price, observable factors that can be used prior to price changes in order to detect price relevant information and the probability, the following trades to be informed.

Third, this attribute of the proposed model can increase the descriptive power of ACD models. The level of how informative a trade is can be measured, using an observable variable, and, therefore, the behaviour of informed and uninformed traders can be further examined. Various issues, such as how informed traders approach the market, the information level of certain trades (e.g., OTC transactions) or how uninformed traders learn, can be further investigated using non parametric analysis. Finally, this analysis empirically extends the Carbon market literature by examining the trading patterns and the behaviour of market participants in EU ETS. Given the unique characteristics of the market, a better understanding of the price impact of various trading strategies could be highly relevant to various aspects of market development.

Drawing on De Luca and Zuccolotto (2006) and Hujer and Vuletic (2007), a Smooth-Transition-Mixture of Distributions ACD (STM-ACD) model is proposed. Following the analysis of the previous chapter, the conditional mean specification for duration is chosen to be the linear-ARMA, as, on average, it outperforms its non-linear counterparts. The associated cumulative density function employed, however, is a mixture of Weibull distributions, where the shape parameter is allowed to vary across three different regimes of trading intensity. The transition between regimes is allowed to be smooth according to Teräsvirta’s smooth transition function.

This modelling allows for various interpretations. The hazard function of a Weibull distribution can exhibit a flat, an upward or a downward slope, depending on the value that the shape parameter takes on.[132] Following the literature mentioned above, the shape of the hazard can indicate the type of market participants. The number of regimes chosen, i.e., three, accounts for non-discretionary liquidity, discretionary/fundamental and informed traders. Furthermore, trading intensity is used as a proxy for measuring the strength of order flow signals, which might be followed by price changes. In addition, the smooth transition function that operates between regimes would account for different paces of learning across participant groups.

[132] Weibull nests Exponential distribution as a special case, when the shape parameter equals one. Consequently the associated hazard function is flat and according to the literature can summarize uninformed trading activity. On the contrary, when the shape parameter is larger (lower) than one, the hazard function follows an upward (downward) slope, different from flat, which is associated with information.

Furthermore, in order to account for a potentially non-linear impact of these types of traders on the conditional mean, a Threshold-ACD (T-ACD) model (Zhang et al., 2001) is estimated. After identifying different regimes of trading intensity that have a statistically significant effect on the density function employed, it might be too restrictive to assume that a simple linear-ARMA specification is sufficient in summarizing the DGP of duration. Therefore, a further dissection of the conditional mean specification according to the previously identified regimes might reveal potential non-linearities. This model is further extended to account for OTC transactions. Then, the in- and out-of-sample performance of the ACD models proposed in this chapter will be compared with the models in Chapter 4. Finally, for robustness, a Burr-ACD model with more flexible hazard functions is estimated to provide further evidence for the trading behaviour of market participants.

One of the most significant findings in the analysis presented in this chapter is that trading intensity is indeed capable of differentiating the trading behaviours of different participants. According to the empirical results discussed in Section 6.3, large or frequent transactions seem to carry information and trading intensity appears to be positively related to this. This is consistent with the theoretical predictions of Easley and O’Hara (1992) and the empirical findings of Dufour and Engle (2000b) that higher trading activity carries more information. Duration is the main determinant but volume contributes considerably, and their effect is greater at the opening and closing of the market, especially during the end of Phase I, when banking was restricted. Moreover, a gradual adjustment that accounts for different speeds of learning seems to be particularly relevant.

In addition, OTC transactions appear to be motivated by information reasons, which can mainly be attributed to their large transaction size. In addition, the non-linear specifications that account for these transactions outperform all other ACD models employed, including the models presented in Chapter 4 (Appendix 4.C). This confirms the empirical findings of the previous chapter, where OTC transactions are found to increase the expected duration. This might be caused either by their large size, which consumes current liquidity, or by their informational content, which might make other traders reluctant to trade. The findings of this analysis provide supportive evidence for the second proposition. Moreover, following the discussion in Section 5.3, informed traders seem to act strategically, as described by Kyle (1985). Consistently, longer durations are associated with no news (Easley and O’Hara, 1992), while the theoretical predictions of Diamond and Verrecchia (1987) do not seem to be relevant in the Carbon market.

The discussion above is particularly important and relevant to various aspects of trading practice in the European Carbon market. First, it highlights the benefit of possessing information before it is fully incorporated into prices, recognising that there are non-pricing variables that can reveal information. On microstructure level, market imperfections are exploitable, even for a very short period of time, and that increases the intrinsic value of acquiring information in "real time". This can affect the attitude of market participants towards the timing and cost of acquiring information.

Second, an increased ability of market participants to indentify informed trades by simply observing past transactions, can increase their market power. Given the "buyer" character of the market, they should easily compensate for potential losses by trading with better informed traders. Their main concern should be liquidity, and therefore spreads should only increase when they observe increased information-based trading and insufficient liquidity. In EU ETS there are significant information and liquidity shocks, as well as strong seasonalities related compliance and a more accurate pricing policy could be proven beneficial. Overall, better risk management, in the sense of information inflow, would lower spreads, and this would contribute to increased market efficiency.

Third, the proposed model can be used for monitoring purposes by regulatory authorities. By identifying informed trading, further action can be taken to protect the market from "manipulation". This is particularly relevant to Viswanathan (2010) who argues that in order for EU ETS to achieve its primary goal, which is emissions' reduction, a fine balance between market innovation and liquidity needs to be found. Given that the model can identify informed trading on a transaction level, regulatory framework can be practically improved by applying real time monitoring tools that can adjust trading intensity and its price impact.

Finally, the remainder of this chapter is organized as follows. The following two sections present the methodology and the empirical findings, respectively. The final section provides a brief summary.

5.2 Methodology

5.2.1 Microstructure

Following the microstructure literature, a typical set-up of strategic trade behaviour is employed. Trades are motivated by the arrival of new private information and reaction to it varies across a pool of heterogeneous market participants that differ in their sensitivity to information and their speed of accumulation of price-relevant knowledge. The estimated model distinguishes between different types of trades, which in many cases determine the associate type of traders. The variables used to proxy this behaviour are the time (i.e., duration) and the size (i.e., number of contracts) of the transaction.

Three major behavioural trading patterns are assumed to be present in the market. First, the informed act solely on the receipt of exogenous private information. This is consistent with Easley and O'Hara (1992) where informed trading can be only understood under the presence of unresolved information. Otherwise, there is no incentive for these traders to enter the market other than liquidity or portfolio reasons that motivate trades unrelated to external information. Considering the time dimension, discussed in the previous section, these traders are the first to receive this information and act in the market. They do so either immediately in large volume trades, as in Glosten and Milgrom (1985), or strategically in segmented patterns, as in Kyle (1985). Easley and O'Hara (1992) indicate that the presence of such traders can be detected through episodes of high trading activity. Therefore, their trading activities could potentially be captured by either large transactions (i.e., large transaction size, independently of whether the associated duration is short or long) or high trading frequency (i.e., short durations, independently of whether the associated transaction size is small or large) or both.

The second type of trades is uninformed. These trades arrive at the market in a random fashion, as described by Easley and O'Hara (1992) and Hujer and Vuletic (2007), and are initiated by the 'non-discretionary liquidity' traders of Admati and Pfleiderer (1988, 1989). Such traders are assumed to transact mainly for liquidity reasons, independent of information. In the context of the Carbon market they are companies that need Carbon allowances purely for compliance reasons. They do not observe order flow and do not extract any price-relevant information out of it. They only follow price resolved information when they need to, and after it goes public.

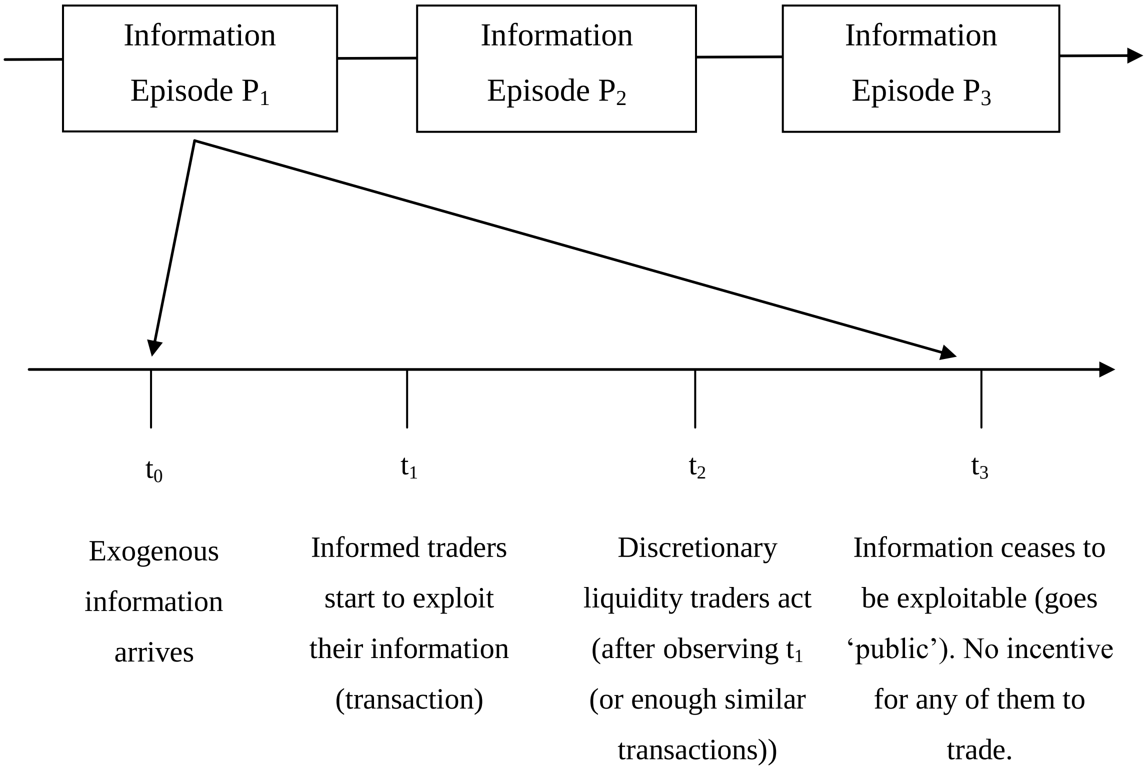
The third type of trades is also uninformed with respect to private information, but is carried out by fundamental traders who have discretion in varying the size and timing of their trades according to endogenous information they can extract from the market. These are similar to the discretionary liquidity traders of Admati and Pfliederer (1988, 1989). They can continuously update themselves analyzing past trading behaviour as revealed by order flow variations, measured by transaction size and trading frequency. They are assumed to act in a similar manner as assumed by PIN models (e.g., Easley et al., 1996) trying to identify informed trading. They are particularly interested in this type of trades either because they are market makers and need to set regret free prices, or because they want to reveal any unresolved information in time, in order to exploit it before it goes public.[133] In addition, similar to Gerhard and Hautsch (2007), they are assumed to have, or acquire, portfolio positions that are revised from desired positions only when information of a certain size or type accumulates over a certain horizon beyond a threshold level. This implies a delayed reaction in time, conditional on that of past actions of other traders, especially that of the informed. Obviously, the quality and the speed of their learning may vary (e.g., Vives, 1993; Jun and Vives, 2004), depending on how they assess the acquired information.

Furthermore, exogenous information is assumed to arrive randomly according to a Poisson distribution. It is first observed by informed traders and then it is progressively revealed by their actions. Fundamental traders capture information signals extracted by trading activity and seek to resolve whether it is sufficiently strong to indicate a subsequent price change. In case they believe so, they act accordingly, buying (selling) the asset upon the arrival of good (bad) news. Finally, when information becomes public, partially because it is revealed by the actions of all previous traders, there is no incentive to further exploit it.

Consequently, information is considered to arrive at the market randomly, creating disturbances in trading activity until it is disclosed in prices. The length and the strength of every information episode depend on the quality of the market, as it can be measured

[133] The underlying assumption here is that both types of traders need to manage a portfolio of some kind. Market makers have a basket of assets that they need to preserve on an optimal level, in order to avoid excessive carrying costs or loss of sales. On the other hand, portfolio managers have a basket of assets aiming at maximizing profit without undertaking excessive risk levels. Although they face different risks, both need to react to new information, especially when the signals are strong enough to substantially change the risk-return relation in their portfolios. In addition, in several studies, such as in Madhavan et al. (1997), when limit orders are allowed, their submission could be considered as market making on one side of the spread. Therefore, these traders could be considered as bearing similar risk to those of market makers.

by variables, such as the depth of the order book or investors' aggressiveness, discussed in Parlour (1998) and Foucault (1999).The following diagram depicts the sequence of trading actions through every information episode. The time line has been expanded for clarity but the informed trades at t_1 can occur at t_0 and the discretionary trades at t_2 can occur at t_3.



The above description of trade types and trader categories describes sequences of trade decisions, for every information episode, that have direct implications on the shape of the hazard function of durations, or else the probability of a trade to occur over a time interval conditional on no trades having occurred before the beginning of this interval. Uninformed traders, following the literature, are assumed to be unaffected by the arrival of new information and, since they are assumed to be naive, they do not change their trading behaviour within each information episode, at least not till the information becomes public. Therefore, the probability of uninformed transactions to occur at every point in time should not change for any length of time. Thus, a flat hazard function, which is a statistical property of the Exponential distribution, should be sufficient in describing their trading.

However, this is not valid for informed traders. They are the first to observe the new information and they have the time advantage to exploit it before it goes public. In case they have the monopoly and there is no time limit and they can choose the time and size of their transactions. However, usually there are more traders, not just a single one, that observe the incoming information and there are also time constraints, e.g., public announcement of the relative piece of information, before it is incorporated into price. In this case, informed traders compete with each other and they need to act within a certain period of time $t_1 < t_3$. In addition, considering that they might be revealed by their trading, they have an even stronger incentive to act closer to t_0, when the episode starts. Consequently, the rate of their transactions, which is relevant only when there is information, should be at its highest as soon as they observe any news. This rate should then decrease to zero as t_3 is approached, since, when information goes public, the probability of observing an informed trade is essentially zero. The conditional intensity for informed trading, i.e., the hazard function $\lambda_{inf}(t)$, can be written as:

$$\lambda_{inf}(t) = \lim_{\Delta \to 0} \frac{1}{\Delta} Prob(t = t_0 + \Delta | P_1; t_1 < t_3) \tag{5.1}$$

where the probability of an informed transaction to occur in time t $(t_0 + \Delta)$, given that there is new information, in a new information episode P_1, and that this information is not public yet $(t_1 < t_3)$, is a decreasing function of time.

In contrast, the fundamental or discretionary liquidity trades/traders have a lagged behaviour that is time dependent and conditional on the order flow created by informed traders. They are present in the market and they continuously update themselves, trying to extract information signal from order flow variations, probably having different learning speeds. After they observe one, they start revealing the informational advantage of informed traders and they have the incentive to exploit it as much as they can before it goes public. According to the previous diagram, this is from t_2 to t_3. However, that signal needs to be sufficiently strong in order to make them deviate from their desired portfolio position. This gives rise to information aggregation that implies an increasing hazard function. This is because it should be more likely for a transaction to occur after some time, at least t_1, has passed (e.g., Gerhard and Hautsch, 2007; Hujer and Vuletic, 2007).

With these depictions of association between the type of traders and the shape of the hazard function one can characterize trades into “regimes” using the threshold auto-regressive methodology of Smooth Transition ACD models. The smooth transition accounts for variations in information aggregation or learning speed.

5.2.2 Smooth-Transition-MixtureACD (STM-ACD)

The model proposed in this chapter aims at capturing the trading patterns of these three categories of traders. It is an ACD model where the associated distribution is a Smooth Transition Mixture (STM-ACD) of Weibull distributions. More specifically, durations are assumed to be divided into three regimes, each one following a differently shaped Weibull distribution. The different shapes are determined by the different values of the shape parameter, which is formulated to be a smooth transition function of a threshold variable. The threshold variable that is correlated with trading behaviour, and that measures changes in order flow, is conjectured to be trading intensity. The underlying assumption is that traders gain information by interpreting order flow signals. Therefore, there are three regimes of trading intensity that correspond to three different regimes of durations. According to the shape of the associated hazards these duration regimes identify the different types of traders.

Then the model can be formulated as follows. Let $d_i = t_i - t_{i-1}$ denote the (raw) duration of transaction i, which is the time elapsed since the preceding transaction $i - 1$. Let x_i denote the diurnally adjusted duration, ψ_i the expected x_i, ε_ι an error term, and S_i a threshold variable that can take up values equal to s_k, where $k = 1{,}2$ is the number of thresholds that determine the three regimes that are assumed to exist. The STM-ACD model can be written as follows:

$$x_i = \psi_i \varepsilon_{i,} \tag{5.2}$$

$$\varepsilon_i | S_i \sim i.i.d., \tag{5.3}$$

$$with\, E(\varepsilon_i | S_i) = E(\varepsilon_i) = 1, \tag{5.4}$$

$$\psi_i = \omega + \sum_{j=1}^{m} a_j\, x_{i-j} + \sum_{j=1}^{q} \beta_i \psi_i. \tag{5.5}$$

The density function of the disturbances is:

$$f(\varepsilon_i | S_i; \tau) = {}^{h_i(\tau)}/_{x_i} \left[\frac{x_i \Gamma\left(1 + {}^{1}/_{h_i(\tau)}\right)}{\psi_i} \right]^{h_i(\tau)} exp\left(- \left[\frac{x_i \Gamma\left(1 + {}^{1}/_{h_i(\tau)}\right)}{\psi_i} \right]^{h_i(\tau)} \right) \tag{5.6}$$

where

$$h_i(\tau) = \gamma^{'} = \gamma_1 + (\gamma_2 - \gamma_1)G_1(S_i\colon g_1, s_1) + (\gamma_3 - \gamma_2)G_2(S_i\colon g_2, s_2), \quad (5.7)$$

$$G_k(S_i\colon g_k, s_k) = (1 + exp\{-g_k(S_i - s_k)\})^{-1}, \quad (5.8)$$

the shape parameter $\gamma' = h(S_i\colon \tau) = h(\tau)$, of the Weibull distribution is a function, h, of the threshold variable, S_i, and a vector of parameter coefficients $\tau = (\gamma_1, \gamma_2, \gamma_3, g_1$ $g_2, s_1, s_2)$, where $\gamma_\mu, \mu = 1, 2, 3$ are the shape parameters of the Weibull distributions in the respective regimes, determined by the threshold variable S_i and $g_k, k = 1,2$ is a vector of the smoothness parameter. For every duration the total shape parameter of the Weibull distribution (γ') is the weighted average of γ_1, γ_2 and γ_3. The weights are determined by two smooth transition functions, G_1 and G_2.

The range of values that the shape parameters can take on has direct implications on the hazard rate and consequently the type of dominant trader present in the market.

- When $\gamma_\mu = 1$ the associated distribution is Exponential and the associated hazard function is flat. This indicates a constant rate of arrival of trades, which coincides with uninformed trading.
- When $\gamma_\mu < 1$ the Weibull distribution has a monotonically decreasing slope, and so is the associated hazard function. This means that the probability of occurrence is higher for shorter durations, which is consistent with informed trading.
- When $\gamma_\mu > 1$ the Weibull distribution is bell shaped and the hazard function has an upward trend. The probability of occurrence increases with time and therefore it could be associated with “fundamental”, discretionary trading.

A Wald test is used to examine the probability of $\gamma_i = 1$.

In addition, the magnitude of g_k describes the transition rate from one regime to another. The higher (lower) it is, the sharper (smoother) the transition. Practically, the transition between regimes could be connected with either “hybrid” trades, such as trades that do not clearly belong to one of the assumed groups of market participants, or with the speed of learning of different traders, capturing how fast can one group, or traders within the same group, learn from information signals revealed by the actions of another group, or other traders within the same group. Consequently, a smoother transition is an indication of slow information resolution or slow pace of learning, probably due to market conditions, such as increased liquidity, that would allow informed traders to hide their actions for longer. In contrast, higher values could indicate improved market efficiency or faster learning.

The threshold variable used is a natural measure of lagged trading intensity, which increases (decreases) when either transaction size is large (small) or duration is short (long):

$$S_i = \frac{v_{i-j}}{d_{i-j}}, \tag{5.9}$$

where v_i is volume of transaction i and j is restricted to 1.

This model combines methodological approaches discussed in three different papers, trying to offer an effective compromise in identifying informed trading. In more detail, first, Hujer and Vuletic (2007) employ the Markov-Switching framework in order to distinguish between informed and uninformed traders, using as reference point the different stages a latent variable can be in. This unobservable variable does not allow the data to determine the associated distribution to each regime. In the present study, similar to De Luca and Zuccolotto (2006), the threshold variable is assumed to be trading intensity, and thus observable, while each regime's distribution is allowed to be data-driven. This is a trade-off between restriction and flexibility. Restricting the choice of the threshold variable might be a strong assumption, but it allows the data to determine the type of trades included in every regime. Consequently, if the level of trading intensity of subsequent trades can be forecasted, it can also be associated with the type of trades. This way the probability of the next trade being informed can also be computed. This is particularly relevant to next chapter's analysis. Along the same lines, Hujer and Vuletic (2007) allow the probability of a trade to be in a specific regime to vary over time, while the probability of transitions is kept constant. In this study, the transition from one regime to the other depends on trading history, and thus it can vary over time, while the thresholds are kept constant, and thus the probability of a transaction belonging to a regime is on average constant. However, the smooth transition relaxes a bit the compromise that trades of the same transaction size and trading frequency will always belong to the same regime. It also opposes the assumption of a "discrete" mixture of distributions, allowing for "hybrid" trades or behaviour indeterminate between trade types. For example, fast acting fundamental traders can be considered as informed by slower acting fundamentals, who might have a slower learning, herd behaviour.

Moreover, the methodological approach employed here is a variation of De Luca and Zuccolotto (2006). They allow the shape parameter of a Pareto II distribution to vary smoothly across two regimes of trading intensity. However, there are some major differences. First, they recognize only two regimes, while here three are employed in order to account for different types of trades. Second, the choice of distribution (i.e., Weibull against Pareto II) constitutes a significant alteration. The three regimes can define different nesting distributions, such as the exponential when $\gamma = 1$, and not merely different shapes of the same distribution. Consequently, the hazard functions have totally different shapes. This enables the identification of distinctly different trading behaviour, rather than just degrees of the same behaviour.

This model is also different from that of Gerhard and Hautsch (2007) in the following. First, it recognises that order flow signals can reveal exogenous information and not only price signals. This is particularly relevant to fundamental traders who are uninformed with regard to the possession of private information, and update their knowledge by observing the intermittent engagement of informed traders. Second, Fundamental traders cannot be distinguished from technical traders, since both need to manage a portfolio of some kind, even if they own only cash to use for transactions, which can be considered as a limited case, and use "Chartism" in order to extract informational content from market signals. The latter are simply considered to be faster in accumulating new information. Along the same lines, the present model recognizes two broad categories of traders, according to exogenous information: informed and uninformed. Uninformed traders can be further divided into liquidity and fundamental traders.

Furthermore, after deriving the threshold values, a T-ACD model, similar to that of Zhang et al. (2001), is employed in order to examine the potential asymmetric effects of trading intensity on the conditional mean. In addition, the existing regimes are dissected even further to account for OTC transaction effects.

The T-ACD model can be written as:

$$\psi_i = \omega_{kc} + \sum_{j=1}^{m} a_{j,c}^{k} x_{i-j} + \sum_{j=1}^{q} \beta_{j,c}^{k} \psi_{i-j}, \tag{5.10}$$

where

$$k = \begin{cases} 1 & when \quad S_i < s_1 \\ 2 & when \quad s_1 < S_i < s_2, \\ 3 & when \quad S_i > s_2 \end{cases} \tag{5.11}$$

$$c = \begin{cases} 1 & when\, x_{i-j} \quad is\ a\ normal\ transaction \\ 2 & when\, x_{i-j} \quad is\ an\ OTC\ transaction \end{cases}, \tag{5.12}$$

and the associated density function is:

$$f(\varepsilon_i|S_i;\tau) = {}^{\gamma_k}/_{x_i} \left[\frac{x_i\Gamma\left(1+{}^1/_{\gamma_k}\right)}{\psi_i}\right]^{\gamma_k} exp\left(-\left[\frac{x_i\Gamma\left(1+{}^1/_{\gamma_k}\right)}{\psi_i}\right]^{\gamma_k}\right), \tag{5.13}$$

where

$$\gamma_k = \begin{cases} \gamma_1 & when \quad S_i < s_1 \\ \gamma_2 & when \quad s_1 < S_i < s_2. \\ \gamma_3 & when \quad S_i > s_2 \end{cases} \tag{5.14}$$

Finally, for testing the robustness of the results, a similar T-ACD model is estimated, where the associated distribution is assumed to be Burr, which is a finite mixture of Weibull distributions that nests the Weibull and hence the Exponential as special cases.[134] This allows for more flexible, specifically non-monotonic hazard functions that could potentially add further characteristics to, or simply confirm, the regimes' analysis. This could benefit current analysis in two ways. First, if STM-ACD is sufficient in capturing regime dynamics, the additional parameters of the Burr should provide consistent results in the form of, similarly shaped hazard functions. This would confirm that the results describe market dynamics and are not imposed by model restrictions. Second, the Weibull distribution is a convenient choice because of the straightforward distinction of regimes., in the sense that the magnitude of the shape parameter determines the shape of the hazard function, and because it allows for flexibility and generality without requiring excessive computational power. However, it can only produce monotonic hazard functions, which imposes a rigid set of trade types, ignoring some stylized facts of the market related to trading activity. In contrast, a Burr distribution can produce non-monotonic hazard functions, which could potentially reveal several peculiarities of trading activity.

As has been discussed in the introductory section, market participants in each group might have a different speed of learning or a different threshold in aggregating

[134] For further analysis of distributional properties and model dynamics see Grammig and Maurer (2000).

information. This might cause some traders in one group to move more quickly. Such behaviour might decrease or increase the life of the information or the probability of transaction occurrence. Consequently, a monotonically increasing or decreasing probability might not be sufficient to describe "hybrid" trades (i.e., trades that may not fit only in one category). The, more flexible, hazard functions of the Burr distribution could provide some evidence of these trades. This would also justify the use of smooth transition functions, which when they are statistically significant indicate that the transition is not sharp and that it might capture trades that cannot fit in one category.

The conditional mean specification of the Burr-ACD model is assumed to be the same as previously (the linear-ARMA):

$$\psi_i = \omega_{kc} + \sum_{j=1}^{m} a_j^k x_{i-j} + \sum_{j=1}^{q} \beta_j^k \psi_{i-j}, \tag{5.15}$$

while the cumulative density function of the mixture of the Burr distribution is assumed to be a discrete mixture of Weibull distributions:

$$f\left(\varepsilon_i \middle| \kappa_k; \sigma_k^2\right) = \frac{\kappa_k}{\xi}\left(\frac{\varepsilon_i}{\xi}\right)^{\kappa_k - 1}\left[1 + \sigma_k^2\left(\frac{\varepsilon_i}{\xi}\right)^{\kappa_k}\right]^{-\left(1 + {}^{1}/_{\sigma_k^2}\right)}, \tag{5.16}$$

where
$$\xi = \frac{\left(\sigma_k^2\right)^{1 + {}^{1}/_{\kappa_k}} \Gamma\left(1 + {}^{1}/_{\sigma_k^2}\right)}{\Gamma\left(1 + {}^{1}/_{\kappa_k}\right)\Gamma\left({}^{1}/_{\sigma_k^2} - {}^{1}/_{\kappa_k}\right)}. \tag{5.17}$$

and k and c have been defined in Eqs. (5.11) and (5.12). For $\sigma_k^2 \to 0$ the Burr distribution reduces to Weibull. The special case of Exponential occurs when $\kappa = 1$.[135]

5.2.3 Estimation and Performance

Estimation is carried out in two steps. First, the STM-ACD model is estimated to determine the values of the threshold variable that define the regimes, and the values of the shape parameters, γ_μ, of the mixture of the Weibull distributions that characterise the slope of the hazard function. Consequently, the types of trades that dominate the regimes are defined. Second, following the identification of the threshold levels, s_1 and

[135] In addition, in the special case, where $\sigma^2 = 1$, Burr distribution reduces to Log-Logistic, with density function $\frac{\left({}^{\kappa_k}/_a\right)\left({}^{\varepsilon_i}/_a\right)^{\kappa_k - 1}}{\left[1 + \left({}^{\varepsilon_i}/_a\right)^{\kappa_k}\right]^2}$, where $a = \frac{1}{\Gamma\left(1 - {}^{1}/_{\kappa_k}\right)} > 0$ and $\kappa_k > 0$.

s_2, and the regimes, further tuning is made by estimating various Threshold-ACD models that allow for different linear-ARMA specifications for the conditional mean, one for each regime. This second step further scrutinizes whether more elaborate modelling of the conditional mean is beneficial in terms of model accuracy. In addition, it investigates OTC transaction effects more closely, and it tests modelling robustness, by allowing higher flexibility in conditional intensities.

The estimation method used is maximum likelihood, using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) (1970) algorithm with numerical derivatives. In addition, in order to deal with heteroskedasticity of known form, the models have been estimated using robust errors.[136] The Log-Likelihood functions $L(\Theta)$ maximized depending on the distributions as follows.

For STM-ACD, T-ACD and T-ACD-OTC (for the last two, only the conditional mean parameters are estimated, since the distribution parameters are known from the first step.

$$L(\Theta) = \sum_{i=1}^{N(T)} ln\left(\frac{\gamma'}{x_i}\right) - \gamma' ln\left(\frac{\Gamma\left(1+{}^{1}/_{\gamma'}\right)x_i}{\psi_i}\right) - \left(\frac{\Gamma\left(1+{}^{1}/_{\gamma'}\right)x_i}{\psi_i}\right)^{\gamma'} \tag{5.18}$$

For Burr-TACD

$$L(\Theta) = \sum_{i=1}^{N}\left\{ln\kappa_k - \kappa_k ln(\xi\psi_i) + (\kappa_k - 1)lnx_i - \left({}^{1}/_{\sigma_k^2} + 1\right)ln\left(1 + \sigma_k^2\left(\frac{\varepsilon_i}{\xi}\right)^{\kappa_k}\right)\right\} \tag{5.19}$$

Optimization is carried out subject to $0 < \sigma_k^2 < \kappa_k$.

In line with Chapter 4, comparison between all models' (i.e., the ones presented in Chapter 4 and the ones presented in this chapter) fitting and forecasting is carried out, using Q-stats(15), the Bayesian Information Criterion (BIC), Q-Q plots, and the in- and out-of-sample forecasting accuracy.[137] The relative results are presented in Appendix

[136] The market gained complexity and liquidity gradually. Duration has decreased considerably over the years, while volume has increased. Consequently, the values of trading intensity are consistently higher over the years, introducing heteroskedasticity of known form.

[137] The relative results, although they are discussed in this chapter, are presented in Appendixes 4.A and 4.B. Furthermore, concerning Q-Q polts, for the STM-ACD model, the following transformation of the standardized durations has been employed: $\tilde{\varepsilon}_i = \left[\frac{x_i\Gamma\left(1+{}^{1}/_{h_i(\tau)}\right)}{\psi_i}\right]^{h_i(\tau)}$. According to Dufour and Engle (2000a) $\tilde{\varepsilon}_i$ follow a unit exponential distribution, which is the distribution $\tilde{\varepsilon}_i$ is tested against.

4.C. One-step forecasts are computed directly from the conditional mean specification. However, long-term forecasts are given by the following equations.

For STM-ACD

$$E_i[x_{i+s}] = \omega \frac{1-(\alpha+\beta)^{s-1}}{1-(\alpha+\beta)} + (\alpha+\beta)^{s-1}\psi_{i+1} \tag{5.20}$$

For T-ACD and T-ACD-OTC

$$E_i[x_{i+s}] = \omega_c^k \frac{1-(a_c^k+\beta_c^k)^{s-1}}{1-(a_c^k+\beta_c^k)} + (a_c^k+\beta_c^k)^{s-1}\psi_{i+1} \tag{5.21}$$

5.3 Empirical Results

5.3.1 Estimation

The first question posed in the analysis of Section 5.2, is whether trading intensity can be used to identify different types of traders. Tables 5.1, 5.2, 5.3 and 5.4 for ECX I, ECX II, NP I & NP II, respectively, provide supportive evidence. The first column of these tables presents the estimation results for the STM-ACD model, as in Eqs. (5.2) and (5.3). The first three coefficients refer to the conditional mean specification, as in Eq. (5.5), while the next three are the shape parameters, γ_μ, of the different Weibull distributions in Eq. (5.7). The last four coefficients are the smoothness parameters, g_k, and the threshold values, s_k, as in Eq. (5.8). *t-statistics* are in parentheses. The next column presents the estimation results of the T-ACD model, similar to that by Zhang et al. (2001), while the last two decompose further the conditional mean parameters to account for OTC transactions, as in Eqs. (5.10), (5.11), (5.12), (5.13) and (5.14). ω, α and β are displayed separately for every regime. The next section of the tables presents the log-Likelihood value, *L*, the Bayesian Information Criterion, *BIC*, and the *Kolmogorov-Smirnov-statistic*, *p-values* in parenthesis. The bottom section presents the hypothesis (Wald) tests that $\gamma_\mu = 1$.

An analytic inspection of the estimation results, especially the part referring to the STM-ACD model, confirms that trading intensity could be used to extract non-price-related information signals from order flow. The different shape parameters, γ_μ, of the Weibull distributions, which are associated with different hazard functions of duration, indicate that it is capable of distinguishing different trading patterns in the different regimes. In more detail, when trading intensity is low (i.e., below the first threshold value, s_1, 0.3943 in ECX I, 1.0123 in ECX II, 0.3684 in NP I and 0.5987 in NP II), the shape parameter of the Weibull distribution for the associated durations, γ_1, varies from 0.9877 in ECX I to 1.1226 in NP I. In addition, a Wald test, consistently, cannot reject the initial hypothesis that $\gamma_1 = 1$ as presented in the bottom section entitled H(O) in each table. Therefore, according to the previous analysis, the distribution reduces to Exponential, which has a flat hazard function and can be associated with non-discretionary liquidity uninformed traders.

In contrast, the other two regimes present a completely different picture. When trading intensity is very high (i.e., above the second threshold value, s_2, which has an estimated value of 0.7090 in ECX I, 2.5099 in ECX II, 0.9380 in NP I and 0.9086 in NP II), the shape parameter, γ_3, is considerably lower than 1 (0.3798 in ECX I, 0.6057 in ECX II, 0.4332 in NP I and 0.5768 in NP II). This corresponds to a Weibull distribution with a downward hazard, which according to the analysis in Section 5.2 describes the trading pattern of informed traders. Along the same lines, in the medium range of trading intensity, i.e., when $s_1 < S_i < s_2$, the shape parameter, γ_2, is larger than 1 (4.5947 in ECX I , 4.3937 in ECX II, 3.0960 in NP I and 3.5080 in NP II). This identifies a category of traders, previously defined as discretionary liquidity fundamental traders. A Wald test (with values of 127.46 and 3462.61 in ECX I, 112.90 and 807.60 in ECX II, 56.78 and 110.01 in NP I and 69.49 and 180.04 in NP II for γ_2 and γ_3 respectively), consistently rejects that the last two parameters equal 1.

An initial assumption that the model is built on, is that informed traders transact only when there is new or unresolved information. Otherwise, their trading cannot be characterised as informed. In addition, they have the incentive to act fast, and before their informational advantage expires. Thus, it is more likely for them to transact sooner upon the arrival of new information, and this behaviour would translate into a downward sloping hazard. According to Easley and O'Hara (1992), this results in increased trading intensity in the market, mainly in terms of increased trading volumes. This proposition seems to be confirmed here. Quiet stages of the market, where trading intensity is low ($S_i < s_1$), appear to be dominated by uninformed liquidity trades, since the hazard function of the associated durations is flat ($\gamma_1 \approx 1$). Consequently, their transaction rate is observed to be rather stable during the trading day and it appears that it is not affected by information arrival. In contrast, higher intensity of trading is associated with increased levels of incoming or resolving information. When trading activity reaches its highest levels ($S_i > s_2$), the associated hazard of durations is a decreasing function of time. Trades in this category are more likely to occur closer to the information arrival, and, according to Section 5.2, are associated with informed trading. This confirms that, indeed, excessive trading intensity, either in the form of high transaction size or trading frequency, is associated with presence of information. On middle levels of trading intensity, when $s_1 < S_i < s_2$, $\gamma_2 > 1$, which indicates an upward slope for the hazard function. The associated trades are more likely to occur as information starts being incorporated into prices, and, consequently, these traders appear to follow information.

Moreover, panel A of Figure 5.1, presents the proportion of trades that belongs to each regime (i.e., uninformed, fundamental and informed) across markets and phases. It shows that the majority, over 70 percent, of the observations is included in the uninformed regime. Consequently, the majority of trades seems to be motivated by compliance reasons and not by speculation or informed trading. This percentage increases in Phase II in both the ECX and NP markets, indicating a more efficient market in terms of asymmetric information and how it is incorporated into price. This may also be seen as a sign of success and establishment of the market, considering also the increase in volumes and market participants Phase II as seen in Figure 3.5.

5.3.2 Smooth Transition, Information Resolution and Learning

The smooth transition feature of the model can capture hybrid trades recognizing that traders can learn from each other, but have qualitatively different learning processes and learning speeds. Econometrically, that is translated into different trading intensity ranges, where the associated durations have a mixed hazard function, which can be captured by a changing shape parameter, due to the smooth transition function. A closer inspection of the estimation tables first reveals that three distinct regimes are clearly identified. The threshold values, s_1 and s_2, are always significant, with the distinction between the first, normal, and the second, fundamental, regimes being more statistically significant (*t-statistics* of s_1 and s_2 are 14.16 and 9.65 ECX I, 29.33 and 18.52 in ECX II, 12.61 and 8.02 in NP I and 7.87 and 4.96 in NP II). Second, the smoothness parameters are substantially large and statistically significant (g_1 is 2.3533 (15.20) in ECX I, 2.3632 (22.37) in ECX II, 2.3840 (12.23) in NP I and 1.6797 (3.52) in NP II, while g_2 is 1.8586 (12.84) in ECX I, 3.7037 (20.73) in ECX II, 1.9867 (9.77) in NP I and 3.5042 (3.20) in NP II), which confirms two things that will be presented next.

The first implication of estimates of smoothness parameters is that the transition between regimes is found to be quite sharp (the smoothness parameters g_k are large), although the shape parameter varies considerably. This could indicate increased market sensitivity towards new information.[138] More specifically, since order flow is assumed to contain price-relevant information, market participants can extract it by translating order flow variations into information signals. Consequently, trading patterns, as they are described by the DGP of duration, change according to the level of trading intensity. A sharp transition between regimes indicates a quick adjustment, which can be translated into increased sensitivity or clear tri-chotomous behaviour of market participants. Consequently, when a transaction belongs to one regime, especially the informed one, it appears to send a sound signal to the market that is picked up by sensitive traders who observe transaction history. This increased sensitivity might lead to phenomena such as “herding” towards the wrong direction and market manipulation, as they are described by Kyle and Viswanathan (2008) and Viswanathan (2010).

The second implication of the estimates of the smoothness parameters is that they recognise various ranges of hybrid trades or different speeds of learning for the types of trades/traders assumed. In more detail, in Phase I $\hat{g}_1$ is larger than $\hat{g}_2$. This means that the learning process of non-discretionary liquidity traders is faster compared to the rate at which discretionary traders learn from informed traders. Alternatively, informed traders can effectively conceal their informed trades. Their actions are not revealed immediately, or at least not as quickly as the informational content of the actions of fundamental traders becomes public knowledge and, therefore, they are followed by uninformed traders. Consequently, the transition from the informed regime to the fundamental is smoother than from the fundamental regime to the uninformed. This means that, fundamental traders’ learning speed is slower compared to the speed at which their informational advantage becomes public knowledge. In contrast, in Phase II $\hat{g}_1$ is smaller than $\hat{g}_2$. This means that the transition from the informed to the fundamental regime is sharper compared to the transition from the fundamental to the uninformed regime. The sharper transition indicates that fundamental traders learn faster in Phase II and it seems that informed traders cannot keep their private information exploitable for long. This can be seen as a sign of the maturity of the Carbon market in Phase II, where there is greater transparency and more efficient

[138] These findings, although, they cannot be exclusively supportive of this sole proposition, they can be considered as a first sign of heightened sensitivity of traders towards order flow variations. This idea will be further discussed in the following chapter.

information flow, and it is consistent with Vives (1993) and Jun and Vives (2004) propositions, who postulate that the speed of learning varies across market participants.

Furthermore, these findings are confirmed by the analysis based on Burr-TACD models. Table 5.5 presents the estimation results for the Eqs. (5.15), (5.16) and (5.17). The first three columns present the estimation results for Phase I in both ECX and NP, while the next three columns refer to Phase II in both markets. In each phase, the first column consists of the parameters of the conditional mean (i.e., ω, α and β) and the shape parameters of the Burr distribution, k and σ^2, for each of the three regimes (i.e., normal, fundamental and informed). The bottom section in each market reports hypothesis (Wald) testing for the shape parameters.

In the regime of normal trades, parameter $\hat{\sigma}^2$ is always close to zero (0.001 in ECX I, 0.0136 in ECX II, 0.649 in NP I and 0.0135 in NP II), while $\hat{\kappa}$ is close to one (1.0387 in ECX I, 0.9911 in ECX II, 1.1647 in NP I and 0.9965 in NP II). Wald tests fail to reject the initial hypotheses that they $\sigma^2 = 0$ (0.37 in ECX I, 0.14 in ECX II, 2.20 in NP I and 6.25 in NP II) and $k = 1$ (1.27 in ECX I, 0.45 in ECX II, 1.41 in NP I and 4.14 in NP II). For this particular subset, the normal regime, Burr reduces to exponential and as it can be seen in Figure 5.9 the associated hazard function is flat. In contrast, in the fundamental and informed regimes the distribution parameters, k and σ^2, vary significantly and Wald tests reject easily all null hypotheses, while Figure 5.9 shows associated hazard functions that are non monotonic. In the fundamental regime $\hat{\sigma}^2$ varies across markets and phases from 0.3805 in NP I to 1.1326 in ECX II, while $\hat{\kappa}$ is larger than 1 and varies from 1.5937 in ECX I to 3.0461 in NP I, which causes the hazard function to be non-monotonically increasing (i.e., increasing up to a point and then decreasing). Finally, in the informed regime $\hat{\kappa}$ varies from 0.7634 in ECX II to 1.2289 in NP II, while $\hat{\sigma}^2$ varies from 0.0164 in ECX II to 1.3540 in ECX I, which indicates a distribution for duration that has a clearly decreasing hazard function. This confirms previous findings from the STM-ACD model.

Moreover, it is really interesting to investigate further the shape of the estimated hazard functions, especially in the fundamental regime. Considering that in this Burr-TACD estimation the regimes are assumed to be discrete, there is a loss of flexibility accounting for different learning speeds. However, the non-monotonic shape of the hazard function in the second regime indicates that, compared to the informed traders, these traders are indeed slower in observing information, because they need to extract it first, hence, the upward slope. However, the subsequent downward slope in the end indicates that after they gain the information, they have the incentive to exploit it, and they act as informed to the remaining uninformed traders. Therefore their trading patterns resemble those of the informed and the hazard function presents a downward slope.

5.3.3 Informed Trading

Intraday Behaviour

The analysis of the estimation results in tables 5.1 to 5.5 confirms the existence of three different regimes corresponding to three different types of trades. The following analysis focuses on how the associated traders behave in the Carbon market. Panel A in Figure 5.1 shows the proportion of transactions in each regime across markets and phases. As expected, the first regime, the normal, includes the vast majority of transactions, while informed traders account for a substantial proportion of trading activity. Fundamental trades account for a maximum of 12 percent in NP I. Another, important outcome is that normal trades have an increased proportion in Phase II, while informed trading decreases in both markets. Again this is a sign of market maturity, consistent with the analysis of the smoothness parameters' results in tables 5.1 to 5.4.

In addition, panel B emphasizes the Probability of Informed Trading (PIN) over the sampled period.[139] It increases throughout Phase I, with a peak in 2007, but decreases in Phase II. This coincides with a general trend of increased demand for futures contracts, as it can be seen in Figure 3.3. Banking (i.e., storage of allowances) between Phase I and Phase II was not allowed and investors started trading intensively in futures contracts (e.g., Daskalakis et al., 2009). This was a very volatile stage of the market and that appears to have been a fertile ground for informed activity (please refer to Figure 3.5 and notice the sharp price changes after June 2006 in both markets).

Moreover, panel C shows informed traders' intraday activity, dividing the trading day into three periods: opening (i.e., the first trading hour), main period, and closing (i.e., the last trading hour). Informed traders appear to be more active, even marginally, in the opening and the closing periods. They are more active before the official closing of the

[139] Unlike previous studies, PIN here measures the proportion of informed trades. Trading intensity is able to identify whether a trade is informed or uninformed. The proportion of informed trades over a period of time could act as an expectation for the future market conditions concerning the level of informed trading. Therefore, PIN in this study is computed as the ratio of informed transactions over the total number of transactions over a period of time.

market.[140] The difference seems to be more obvious in NP. This could probably be explained by the fact that NP closes earlier and that OTC EUA holders can enter the market directly. In addition, the standardized contracts allow traders to enter both markets, and traders that possess information might choose the appropriate environment, such as the more volatile session before the official closing, for their trades.

Strategic versus Sequential Trading and Timing Advantage

Furthermore, the STM-ACD model, and especially the interpretation of the hazard functions of durations, allows for the propositions of Glosten and Milgrom (1985) and Kyle (1985), concerning whether informed traders act at once or strategically, to be tested. Panel A of Figure 5.2 shows the value of the coefficient gamma across trading intensity. This graphically presents the estimates of γ_μ from Tables 5.1, 5.2, 5.3 and 5.4. When trading intensity is low $\gamma_\mu = \gamma_1$ is always lower than one, when trading intensity is in a middle range $\gamma_\mu = \gamma_2$ is higher than one, while when trading intensity is high $\gamma_\mu = \gamma_3$ is consistently lower than 1. This indicates that informed trading, which is associated with a downward slopping hazard (i.e $\gamma_\mu < 1$), is related to either high trading volume, which would be supportive of Glosten and Milgrom (1985), or high trading frequency, which would be supportive of strategic trading according to Kyle's (1985) propositions. This finding does not provide a specific answer to whether informed traders act according to Kyle (1985) or according to Glosten and Milgrom (1985), but it confirms Easley and O'Hara (1992) who argue that informed trading is associated with increased trading activity.

In addition, Table 5.6 reports the basic statistics of duration in the first four columns, of volume in the middle four columns and of trading intensity in the last four columns, in all markets and phases, focusing on each year separately. The statistics are dissected for Buyer (B) and Seller (O) initiated transactions across the three regimes, namely normal, fundamental and informed. Panel B of Figure 5.2 and Figure 5.3 draw from Table 5.6 and present the relative change of volume and duration across regimes. More specifically, panel B of Figure 5.2 presents graphically the average transaction size across market and phases in each of the three regimes, while Figure 5.3 shows the average duration of Buys (B) and Sells (O) in all markets and phases. Both Figures

[140] This is really relevant to the Bid-Ask spread analysis presented in the following chapter, as it has a significant impact on implied spreads and price change volatility.

confirm the findings presented in panel A of Figure 5.2 and show that informed traders act either fast (e.g., Figure 5.3 shows that the average duration of informed transactions is shorter compared to the other two) or in large quantities (e.g., panel A of Figure 5.2 shows that the average transaction size is higher in the informed regime).

In contrast, trades in the normal regime are characterised by longer durations and low volumes. As trading intensity increases, traders appear to be more connected to information. In more detail, the large difference in durations between the normal and the other two regimes, explains partially the higher t-statistics of the first threshold value s_1 in Tables 5.1 to 5.4. Furthermore, the difference in volume between the normal and the other two regimes is not that large, excluding ECX II. This leads to a primary conclusion that the main difference of regimes is due to the large differences in the speed of trading (Figure 5.3). This is consistent with Kyle (1985).

Moreover, panel C of Figure 5.2 presents the autocorrelation of order flow in each regime across markets and phases. The autocorrelation of Buyer and Seller initiated transactions is further examined. If it is higher in the informed regime, that would be an indication of strategic trading. The higher autocorrelation of order flow in the informed regime is supportive of Kyle (1985) because it indicates that many trades move towards the same direction, which might be part of strategic trading. Indeed, it is everywhere higher, apart from ECX II, which consistently appears to be structurally different. In addition, the fundamental regime once again appears to be very different from the other two regimes, which strengthens the validity of the estimation results, providing empirical evidence for the existence of three distinct regimes.

Furthermore, Table 5.7, which complements Figure 5.2, examines further whether trades in the informed regime are strategic. The first horizontal section of Figure 5.7 reports the proportion of identical trades, as they are compared to the previous transaction, in each regime across markets and phases. Identical trades, especially when the duration is short and the volume low, would indicate strategic behaviour in the form of segmented trades. The second section of the Table 5.7 presents the proportion of identical trades in each regime across markets and phases, as they are compared to the previous transaction of the same group. This section tries to capture cases, where there are other, probably uninformed, transactions between segmented trades. The final section reports the average duration of identical transactions and unique transactions in all markets and phases.

A close inspection of Table 5.7 reveals that the proportion of identical trades, when compared with the previous transaction, is consistently higher in the informed regime. This practically means that 31.77 percent in ECX I, 19.51 percent in ECX II and 29.72 percent in NP I of the transactions are identical to the previous one, which is a strong indication of strategic behaviour. In contrast, no identical transaction is found in NP I, which means that Kyle's (1985) propositions cannot be confirmed. The same picture is confirmed when the current transaction is compared with the previous one from the same regime, but only in ECX I and NP II. ECX II, once again, appears to be different. Furthermore, the last section shows that the average duration of the identical transactions, compared to the previous trade, is consistently shorter in the informed regime. This is another sign of small segmented trades, and, thus, of strategic behaviour. Consequently, the above tables and graphs cannot provide a clear picture of the behaviour of informed traders. However, they appear to act strategically in ECX I and NP II, while in ECX II they seem to act at once, probably because their actions can be revealed more easily, which is consistent with the analysis of smoothness parameters in Tables 5.1 to 5.4.

Another important observation can be made from the previous findings presented in Table 5.7 and panel A of Figure 5.2. It seems that there is a large proportion of trades with informational content that is not apparent in price, at least not contemporaneously. More specifically, panel A of Figure 5.2 shows that in ECX I almost 20 percent of the trades can be characterised as informed, while, according to Table 5.7, 30 percent of these trades is not characterised by a change in price. In Table 5.7, the proportion of identical transactions, compared with the previous one, of the informed regime is 30 percent. This strengthens the belief that exogenous information can be revealed by other non-price-related variables of order flow, such as trading intensity, at least for some of this 30 percent of informed trades. Consistent with previous literature, such as the inventory and information models, trading intensity in the Carbon market appears to carry price-relevant, unresolved information that is expected to be incorporated favourably into prices.

Informed Trading and OTC Transactions

Furthermore, an issue already introduced is the relation between OTC transactions and informed trading. Panel A of Figure 5.4 presents the proportion of OTC in each regime, across markets and phases, and it shows clearly that the proportion of OTC transactions in the informed and the fundamental trading regimes is remarkably higher than in the normal regime, especially in ECX. In addition, panel B of Figure 5.4 shows that the average volume per transaction of OTC trades far exceeds that of non-OTC transactions. This is more pronounced in ECX II. This seems to be consistent with a previous comment, that in ECX II informed traders seem to act at once, hence the larger average volume. Panels A and B of Figure 5.4 seem also to confirm stylized facts reported in the literature on other markets concerning the informational content of OTC trades. OTC trades appear to be highly correlated with informed trading. In addition, looking at the estimation results in Tables 5.1, 5.2, 5.3 and 5.4 the *t-statistics* of the OTC-relevant coefficients in the informed regime are consistently higher than those of non-OTC transactions.

Good and Bad News

In addition, the formulation of the STM-ACD model allows for the discussion of the propositions of Diamond and Verrecchia (1987) and Easley and O'Hara (1992) concerning exogenous information and duration length. The first paper shows that longer durations are associated with bad news, while the second paper interprets longer durations as absence of new and price-related information. Considering that informed traders would buy (sell) upon the arrival of "good" ("bad") news, consistently longer durations for Seller initiated transactions would provide supportive evidence of Diamond and Verrecchia's (1987) propositions. In contrast, longer average durations in the normal regime, which is less related to information, would be consistent with the propositions of Easley and O'Hara (1992).

Panels A, B, C and D of Figure 5.3 present the average duration in each regime, dissected into Buyer (B) and Seller (O) initiated trades, in ECX I, ECX II, NP I and NP II, respectively. They show first, that the average duration in the normal regime, for both Buys and Sells exceeds by far the average duration in the other two regimes in both markets and phases (e.g., in ECX I the average duration of Buys and Sells is 582.82 and 419.14 in normal regime, 56.09 and 27.07 in fundamental regime and 20.66 and 7.73 seconds in the informed regime, respectively). Following the initial assumption that the normal regime captures non-discretionary liquidity traders, and thus non-information-related trades, while the other two regimes are information related, these figures tend to confirm Easley and O'Hara's (1992) proposition that no, or delayed, trades are related to lack of information. In addition, taking into account that the normal regime accounts for more than 60 percent of total trading, as panel A of Figure 5.1 reports, the Carbon market could be described as a pool of uninformed, non-discretionary, liquidity transactions with episodes of exogenous information, that triggers fast, information-related, trading.

Focusing on the average duration of Buys and Sells, Figure 5.3 fails to provide supportive evidence of Diamond and Verrechia's (1987) propositions. The average duration of the transactions in the informed regime is of particular importance, because it is an indication of how informed traders act upon the arrival of "good" or "bad" news. The average duration of Sells (O) is consistently shorter in all markets and phases. Especially in the informed regime, Buys and Sells are 20.66 and 7.73 in ECX I, 4.37 and 2.48 in ECX II, 36.53 and 33.31 in NP II and 32.01 and 27.57 seconds in NP II, respectively. Contrary to Diamond and Verrechia (1987), longer durations are associated with "no" price-relevant information and not with "bad" news, since informed traders seem to react faster upon the arrival of bad news than upon the arrival of "good" news. However, the shorter duration of Sells might have another explanation. Panels A, B and C of Figure 5.5 show the number of Buyer (B) and Seller (O) initiated transactions over the years in both markets. In more detail, panels A and B report the absolute numbers of Buys and Sells in ECX and NP, respectively, while panel C reports their difference in both markets. All these graphs indicate that the Carbon market is predominantly a "Buyer" initiated market with an increasing demand over the years. This explains the shorter durations of Sells, since a Seller initiated transaction is more likely to find a Buy to match.

Trade Informational Content

The estimation results also shed light on the propositions of Pascual et al. (2004), who argue that trading frequency increases after a trade with high information content, and therefore, price discovery improves after such trades. Table 5.8 presents the average duration, volume and trading intensity before and after an informed trade (as identified by the STM-ACD). Trading intensity increases after an informed trade in both markets and phases (before and after; 1.4544 and 1.4861 in ECX I, 1.4232 and 1.4799 in ECX II, 1.5232 and 1.5754 in NP I and 1.7372 and 1.9819 in NP II). However, this increase is caused mainly by shorter durations (before and after; 0.6708 and 0.6613 in ECX I, 0.7851 and 0.7678 in ECX II, 0.6441 and 0.6172 in NP I and 0.7995 and 0.7807 in NP II) and not by higher volume (11.8333 and 11.6745 in ECX I, 10.3196 and 10.079 in

ECX II, 10.3582 and 10.6613 in NP I and 10.8739 and 10.6502 in NP II). This is consistent with Pascual et al. (2004). It also strengthens previous findings too, since market participants, namely the fundamentals, who observe order flow, appear to be sensitive to information and, when an informed transaction arrives at the market, they tend to follow it. This results in increased trading frequency, which is higher than in normal regime (trading intensity in normal and fundamental regimes is 0.8426 and 1.0333 in ECX I, 0.9301 and 1.1932 in ECX II, 0.8888 and 1.1835 in NP I and 0.8501 and 1.1734 in NP II, respectively).

5.3.4 Model Application on a Real Event

The application of the model in identifying informed trading can be further tested, by examining a real event in the market. Figure 3.5 presents the daily prices and the aggregate volumes in both markets over the sample period. There are several sharp price changes that are worthy to be discussed, with the sharpest being near market inception (28 April 2006). At the time Spain announces lower emissions than allowances resulting in a sharp price drop. However, since there is a lack of data around this period, it does not receive any particular attention. Another sharp fall is observed after the collapse of Lehman Brothers on 15 October 2008, with Carbon prices falling almost a month later. However, this is mostly related to investors' beliefs and to insecurity caused by the unstable demand for industrial goods, and consequently for allowances. Another price fall occurs on 1 August 2008. Carbon prices follow the sharp downturn in commodities' prices. This was a high volume and volatility environment, as confirmed by the highest unconditional volatility observed in intraday prices for all data sets employed, which might have created a fertile ground for informed trading. This period will be examined in order to test the explanatory power of the model, although it does not represent the more radical of graphical changes, for two reasons. First, the main source of the price change lies in related commodity markets and, second, there is unobstructed availability of data.

Panels A and B of Figure 5.6 present the proportion of the different types of traders (normal, fundamental and informed) around the price event on 1 August 2008 with the focus lying on the time three days before and three days after the event. In both markets informed traders increase their transactions two (30 July 2008 in NP) or three (29 July 2008 in ECX) days prior to the price drop. Their transactions account for around 13 percent in ECX and 8 percent in NP of total trading. The next to follow are the

Fundamental traders, who start trading more intensively one day after (30 July 2008 in ECX and 31 July 2008 in NP) the increased presence of informed traders. Their presence accounts now for around 14 percent in ECX and 6 percent in NP of total trading. According to the assumptions of STM-ACD, this happens because they observe and respond to market signals with a lag. In this example their lag seems to be nearly one or two days. In contrast, normal traders' market share decreases to around 77 percent in ECX and 92 percent in NP. The presence of informed traders reaches its maximum, 14 percent, on 31 July 2008 in ECX and around 11 percent on 1 August 2008 in NP. It is, once again, followed by fundamental traders, who reach their maximum, 16 percent, in ECX on 1 August. Normal traders' presence decreases to a minimum of 72 percent in ECX and 88 percent in NP. They are the last to react and probably do so because of the general market movement.

In order to further examine the actions of informed traders, it would be useful to look at the direction of their trades, Buys (B) or Sells (O). Panels A and B of Figure 5.7 show the difference in Buyer/Seller initiated transactions three days before and three days after 1 August 2008. Both figures, in ECX and in Nord Pool, indicate that informed traders, not only increase their trading, but, unlike all other market participants, also start selling contracts before everyone else. The other two market participants do follow, but with a lag. The difference of Buys minus Sells is still positive for normal traders but it is continuously decreasing. In contrast, on 1 August 2008 fundamental traders follow the informed ones and sell more than they buy. Their actions follow informed traders' and their lag seems to be one or two days long.

However, what seems strange is that immediately after the price drop informed traders start buying again and the other two groups follow, although not with a rush. This can be further explained by examining panels A and B of Figure 5.8. These figures show the aggregated volumes of Buys and Sells, along with the average duration, of each one of the three regimes, dissected into Buys and Sells, three days before and three days after 1 August 2008. Here, informed traders clearly start selling more before the event and they clearly start buying immediately after it. Fundamental traders follow with a two-day lag, at least concerning total volume. In contrast, uninformed traders in the normal regime continue buying, but the total volume of Sells fluctuates according to price (e.g., they sell more after the sharp price drop) and act contrary to informed traders (i.e., buy more when informed sell more). Supposing that informed traders possess price-relevant information, they seem to use it accordingly. On 30 July they increase their selling. In contrast, the associated volume of Buys of normal trades increases as well. This means that the uninformed buy from the informed traders. Exactly the opposite occurs on 1 August. The increased demand of informed traders is met by the increased selling initiated by fundamental and uninformed traders. This is a strong indication, at least in this case, that STM-ACD indeed captures informed traders, who exploit their information to their benefit.

5.3.5 Fitting and Forecasting

Finally, it would be interesting to examine whether the modelling framework adopted in the analysis of this chapter improves the fitting and forecasting of duration. Panel B of figures 4.1, 4.2, 4.3 and 4.4 show that the T-ACD models outrun their counterparts in terms of fitting. In addition, when OTC trades are taken into account, fitting seems to be improved even further and thus the higher ranking. Along the same lines, Tables 4.8, 4.9 and 4.10 confirm the previous findings about fitting. In particular, T-ACD models, with or without OTC transactions, provide far better “in-sample” forecasts and always obtain first place rankings. The STM-ACD comes second best, and has a better fit than its linear counterparts with a sole distribution. T-ACD also provides superior out-of-sample forecasts, indicating the presence of the non-linear effects emphasized in the previous chapter. Moreover, OTC transactions appear to be more relevant to Nord Pool, especially in Phase II. Overall, trading intensity appears to be particularly relevant in the Carbon market and its analysis significantly improves the accuracy of duration models.

5.4 Summary

The analysis of the chapter proposes trading intensity as a measure of information revelation in the market of Carbon Allowances. A Smooth Transition ACD model with a mixture of distributions is used to identify different regimes of duration characterised by different trading intensity. This is conjectured to identify different trading behaviour of market participants. The results make a strong case that private information is indeed revealed in the rate of order flow as measured by trading intensity. Informed traders seem to act first and, at least in some cases, prior to price event occurrence. This is evidence supporting Glosten and Milgrom’s (1985) propositions about the intense activity of informed traders. Informed trading behaviour, however, seems to be followed by similar, but slightly delayed and less intense behaviour of some other traders, identified as fundamental traders. Their delayed order flow prolongs the effect, or revelation, of the information that the order flow of the informed has started, increasing trading frequency. These fundamental traders act as ‘informed’ to the un-informed, who are the third much larger group of traders/trades present in the market, while both seem to follow a learning process. This is consistent with the social learning literature, where traders exhibit different learning speeds. The order flow of this third group is characterised by high duration and lower trading intensity. The behaviour of the “fundamental” traders gives credence to Kyle’s (1985) gradual activity of traders. Accordingly, the results support Back and Baruch (2004) in that there is a time dimension and that different traders trade at different rates.

Amongst OTC transactions there is a higher proportion of informed and fundamental traders. This is evidence supporting the stylized fact in many other markets that informed trades tend to emanate over the counter. This obviously points to the informed traders to trade off the main trading mechanism of the market perhaps due to reasons related to price or trade anonymity or greater depth, lower costs and liquidity. It also points to the fact that the OTC market seems to take the lead in private information revelation through order flow. This piece of information seems to be extremely relevant to duration modelling, since it significantly improves fitting and forecasting.

The Carbon market is a Buyer market where a higher proportion of trades are, on average, Buyer, rather than Seller, initiated. Duration of Buyer initiated transactions is longer than that of Seller initiated transactions. This opposes Diamond and Verrechia’s (1987) proposition that longer durations are associated with bad news. In contrast, uninformed trades seem to be related with longer durations, which, accordingly, are more likely to be associated with no news, instead of bad news.

Chapter 6

Trading Intensity and Intraday Price Discovery

A Dynamic Expectations Joint Model

6. Trading Intensity and Intraday Price Discovery

6.1 Introduction

Empirical market microstructure literature recognizes information and liquidity as the driving forces of intraday price formation. O'Hara (1995) conjectures that, on an intraday level, the market's equilibrium is determined by the trading strategies of the market participants, according to their access to relevant information and their liquidity desires. A market is viewed as a mechanism of information aggregation, in which investors' beliefs about the fundamental value of the underlying asset move prices accordingly. Therefore, information is generally considered to have a permanent impact on prices. However, markets also act as a matching mechanism for instantaneous demand and supply at every particular point in time. Liquidity considerations might force the price to temporarily deviate from its "fair" value. Therefore, liquidity is mainly associated with transitory price effects.

Both information and liquidity are particularly important for the presence of market specialists, especially in a relatively illiquid environment such as the Carbon market (see, inter alia, Benz and Hengelbrock, 2008) where dealers have been introduced to support liquidity. Access to information and marginal levels of liquidity determine dealers' risks and consequently investors' costs mainly in the form of the Bid-Ask spread. Unlike Stigler (1967), who sees trading costs as a sign of market "imperfection", Demsetz (1968) raises the importance of spreads as a natural outcome of trading activity. Market makers provide a constant demand and supply in the market and they need to be rewarded for bearing the risks and the costs of immediate execution of trading. This is the price of "immediacy", which implicitly connects costs with trading activity. Therefore, apart from explicit trading costs, such as market charges, the time dimension and the size of a transaction are seen for the first time as inextricable determinants of quoted prices and, consequently, of costs. Therefore, trading costs can no longer be ignored or be seen as mere market frictions since they are directly related to trading activity.

Along the same lines, Bagehot (1971) emphasizes that, even without explicit costs, or even trading costs, Bid-Ask spreads should still exist. He emphasizes the time dimension of information. At any given point in time, market participants do not all possess the same level of information, and their learning rates might differ.

Consequently, dealers need a compensation for offering liquidity in the presence of better informed traders. Naturally, these two fundamental market forces, information and liquidity, are seen in a substantial proportion of the theoretical and empirical literature as the main determinants of the spread.

Previous studies of the Bid-Ask spread have identified three main components.[141] The first, already mentioned in Demsetz (1971), is associated with per transaction execution costs, and is called the order-processing cost.[142] This cost is largely fixed and therefore should fall, at least per unit, with trading volume (see, inter alia, Huang and Stoll, 1997; Ahn et al., 2002). However, diversified market making in a portfolio of assets could amortize more efficiently this cost and this relation may weaken.[143] In addition in a highly competitive market Bid-Ask spreads should equal the expected marginal cost of providing liquidity, in which case the order-processing cost may be irrelevant (see, inter alia, Bollen et al., 2004).

The second component identified in the literature is the inventory-holding cost.[144] A distinct class of models, "Inventory models", originated by Tinic (1972), Tinic and West (1972) and Benston and Hagerman (1974), recognizes that market specialists face a complex balancing problem between inventories and incoming order flow. Dealers use the price to balance supply and demand trying to accommodate two types of risk. The first is related to their inventory position, in the form of excessive carrying cost or lost sales. High inventories might result in excessive carrying costs, especially overnight, while low inventories, might be insufficient in supporting sharp changes in incoming order flow. The second is related to the "fundamental" price change, in the sense that sharp price changes might change the value of their portfolios. Since inventories operate as a buffer to market activity, these models postulate that the main determinant of price changes, apart from a random walk component associated with information, is order flow deviations. These deviations are, by assumption, unrelated to the future value of

---

141 For an extensive analysis refer to O'Hara (1995), Madhavan (2000), Hasbrouck (2007) and Stoll (2003).

142 It includes items such as the exchange seat, floor space rent, computer costs, informational service costs, labour costs and opportunity costs for the market maker's time (Stoll, 1978). The majority of empirical microstructure models incorporate this cost component as a constant. This coefficient captures a constant fee that is charged by market makers as a compensation for supplying liquidity and it represents their efficient operations, their competitive advantage and their market power (Garman, 1976).

143 More liquid assets could bear a higher proportion of these costs, since they can reach the breakeven point faster. This being analogous to liquidity, amortization still allows for the development of economies of scale, but their effect should be limited, according to the relative weighting.

144 Some researchers (George et al., 1991; Hanousek and Podpiera, 2004) reject inventory-holding as a spread component arguing that this component exists only in the extreme situation of a general trading pressure.

the asset, and their price impact is temporary (e.g., Hasbrouck, 1988, 1991a, 1991b). They distinguish the informational, or permanent, from the liquidity, or transitory, effect of a trade. Therefore, inventory positions affect the quoted spread on the short-term.

Several studies provide different explanations for the price and spread impact of inventory positions. Garman (1976), in a rather simplistic approach, emphasizes the "viability" of dealer's positions. Prices are set to avoid market "failure". Amihud and Mendelson (1980), extending Garman's model, assume that there is an optimal, or preferred, inventory position. Dealers set their prices in order to maintain or return to this position and therefore they are seen as trying to maximize their profit.[145] Stoll (1978) and Ho and Stoll (1981, 1983) relax the implicit assumption of a risk-neutral market maker in the previous model, where the "market power" is the main determinant of the spread, and see the dealer as a market participant trying to maximize the utility of his/her portfolio. In this case, the cost of immediacy is identified as order-processing costs, liquidity considerations and adverse-selection; this last cost component was developed in another class of models that will be discussed in the following paragraphs. Therefore, the main determinant of the spread is the risk aversion of the dealer about two types of risk, the uncertainty in returns and the uncertainty in the incoming order flow, in terms of size and direction.

Furthermore, O'Hara and Oldfield (1986) examine the dynamic pricing of a risk-averse market maker in a competitive environment, where dealers can interact with each other and other market participants who can submit limit orders on only one side of the spread. Consequently, the size of the spread is connected with the type of order (i.e., market or limit order) and the "gravitational pull" (i.e., tendency to submit orders of similar sign and type). Some studies reject the inventory-holding component, arguing that this component exists only in the extreme situation of a general trading pressure, such as extreme liquidity or price volatility conditions.

There are four important points that have been emphasized by inventory models and are relevant to this study. First, these models emphasize the temporary effect of liquidity

[145] In Garman's (1976) model, dealers cannot borrow money and market failure occurs because they run out of cash or inventory. Therefore, they set their prices in order to manage that risk. However, the basic assumption that they set their prices once and that they cannot borrow money are rather restrictive. On the contrary, in Amihud and Mendelson's (1980) model, dealers have an optimal inventory position and adjust prices (i.e., decrease (increase) both Bid and Ask for long (short) positions) in order to maximize profit, instead of minimizing risk. Madhavan and Smidt (1993) and Hasbrouck and Sofianos (1993) provide further empirical evidence on optimal inventory levels. Therefore, even when dealers are at the optimal inventory level, they still quote different Ask and Bid prices, and this spread is the cost for departing from their desired level.

and risk aversion on prices and spreads.[146] At this early stage of the literature, liquidity is measured by the size of inventory positions or by changes in these positions due to transactions.[147] Risk aversion, introduced by Stoll (1978) and Ho and Stoll (1981, 1983) appears to have an unambiguous effect on the Bid-Ask spread and several measures of risk have been proposed in the literature to account for the price impact of various sources of risk. Among them, Demsetz (1968) uses transaction rate, while many studies use return volatility (e.g., Tinic, 1972; Stoll, 1978b; Harris, 1994).[148] Bollen et al. (2004) show a positive relation between volatility and inventory-holding cost. Second, Stoll (1978) shows the importance of trading volume. He also shows that implied spread exhibits a U-shaped curve over order size. The practical implication is that there is an optimal level for transaction size that minimizes costs. Third, order type affects dealers' decisions about setting risk-averse prices (e.g., O'Hara and Oldfield, 1986). This is plausible since limit orders can be seen as market making in one side of the spread and hence the dealers' inventories and investors' portfolios interact in price formation. Fourth, as inventory models become more complex by, for example, including more factors about uncertainty in dealers' balancing problem, the importance of time becomes obvious. The last three comments about transaction size, time and type of order are of particular relevance to this analysis.

The third component of the Bid-Ask spread is the adverse-selection cost. Bagehot (1971) introduces the revolutionary idea that even without explicit or inventory trading costs, spreads would still exist as a natural outcome of the trading process, since Bid and Ask market equilibrium prices would be different due to asymmetric information. He considers dealers as market participants who, sometimes, transact with better informed traders. In this case they lose money, but they can compensate by dealing with

[146] Especially risk aversion has gained considerable attention and in some studies it is considered as a separate cost component (e.g., Bollen, 2004). However, risk aversion cannot be measured directly. It is usually captured by a premium charged on top of a risk free rate that accounts for risk free market conditions, or perfect monopoly. In this study, risk aversion is measured as the dealers' sensitivity towards their expectations about risk, and how long are they expected to be exposed to it. Risk is measured in relation to information, price volatility and liquidity.

[147] Madhavan (2000) mentions that in a dealer market where all transactions are operated by dealers the change in inventory positions is simply the opposite of the signed volume. However, in a hybrid or a pure limit order market, where market participants are allowed to transact with each other without the intermediation of a market specialist, the change in inventories can no longer be measured by signed volume. This does not necessarily mean that limit orders do not contribute to spread formulation. Limit orders can be seen as market making on one side of the spread. They could also be part of a strategy referring to a portfolio of assets, similar to the inventory of a dealer. Although more recent models do not explicitly model inventory costs, they account for liquidity considerations and the transitory price impact of a trade, mainly distinguishing between the permanent and the transitory price impact of time, volume and order flow variations.

[148] Other measures of risk include the idiosyncratic risk of the asset (see Benston and Hagerman, 1974) and the average standardized price difference (see Tinic and West, 1972).

uninformed liquidity traders. Therefore, spreads account for these dealers' risk aversion towards adverse-selection risk. In addition, since the associated information refers to the future value of the asset, price changes due to adverse-selection tend to be permanent. These ideas have been extensively developed by a new class of models that are generally referred to as "information" models. A key aspect of the empirical analysis presented in this chapter is the informational content of trades, and whether these trades can be used to identify the presence of informed traders.[149]

Information models can be divided into two broad categories according to what types of market participant ("players"), are recognized and how they interact ("game"). A common key aspect to both groups is the division of traders according to the level of information they possess and their learning process. This is defined as the degree of their "naivety". First, the sequential trade models, originating in Glosten and Milgrom (1985) and extended by Easley and O'Hara (1987, 1992) and Easley et al. (1997, 2002), recognize two types of traders, informed and uninformed, liquidity traders. Dealers' spreads include a component as a compensation for transacting with informed traders. These models emphasize the permanent informational content of trades. Similar to inventory models, the size and the time of a transaction convey information. This information can potentially reveal the type of a trader.[150] Informed traders are assumed to act at once exploiting, as much as they can, their informational advantage. Therefore, the expected size of a transaction and the probability of the transaction to be informed determine the convergence to fully informative price and the spread.[151] More recently

---

[149] Trading process is seen as a "game" that involves informed traders. The early market microstructure literature, involving theoretical inventory and information models, draws from the "Rational Expectations Equilibrium" (REE) and "Game Theory" literature and tries to explain the trading process and the intraday price formation theoretically. Early studies examine issues like the speed of information dissemination, information aggregation, learning and herding, as well as inventory positions, time, order size and order type. Some models, such as Kyle (1985) and Glosten and Milgrom, (1985), have been very influential and have been empirically confirmed from many aspects. However, since they depend heavily on game theory, they all have a major drawback. In order to make safe conclusions, the "game" (e.g., market type, market participants and order type) must be known. Otherwise, some drastic assumptions must be made. The focus of the present analysis, though, emphasizes the empirical aspects of market microstructure and therefore the use of the REE and game theory references will be limited to the most influential papers. For two excessive and excellent reviews of this literature please refer to Burrnnermeier (2001) and Vives (2008).

[150] The majority of these models use transaction and not limit order data. The market is assumed to be a dealer market and the equilibrium price is determined by the order flow. However, more recent studies, deviate from this setting and empirically examine hybrid markets, such as NYSE. In addition, they allow other variables to determine market equilibriums.

[151] A key aspect of these models is the incoming order flow, which is assumed to follow a Poisson process. Dealers try to anticipate whether the incoming trades are informed and they formulate the Probability of Informed Trading (PIN). A key element of PIN is the expected volume which might be either large or small. Although PIN calculation does not originate in sequential models, it has been used as a model assumption and not as a model prediction. Therefore, they have been developed simultaneously.

Easley et al. (2008) introduce the idea that expected duration is an integral part of PIN. The arrival time of transactions or orders can reveal price-relevant information and can provide an indication of the presence of information. Consequently, the information level of trades can be measured by the expected arrival time.[152]

Although sequential models might be sufficient in explaining after-shock informative prices, their main drawback is that they see the transactions as a sequence of distinct events and they do not allow for development of informed strategies. Full information convergence is assumed to happen immediately after the event. This gave the incentive to the second category of information models; the strategic models, originated by Kyle (1985) and developed by Kyle (1989), Foster and Viswanathan (1993), Holden and Subrahmayam (1992) and Back (1992), to allow for progressive information resolution. These early strategic models similarly recognize two types of traders, informed and uninformed, but informed traders can strategically cover their advantage and thereby exploit it further.[153] However, uninformed traders are still considered naive. In contrast, Admati and Pfleiderer (1988), as well as Foster and Viswanathan (1990) and Spiegel and Subrahmayam (1992), distinguish the uninformed traders as either "discretionary" or "non-discretionary" liquidity traders. This way, they recognise that there is a part of uninformed traders that tries to extract information from past trading they learn from it and they choose the time and size of their trading accordingly.[154]

The present analysis draws the following points from information models. First, a common feature of sequential and strategic models is that they distinguish between types of trades, and consequently types of traders, recognizing that the price impact of their transactions might differ. Second, they see past trades as indicative of the type of

[152] In the present analysis, following the findings in the previous chapter, the information role of transaction size is also examined. This provides the framework for a further analysis of PIN. Employing the same methodology, the expected trading intensity can be forecasted, which is expected to have an impact on PIN. This could further enrich the analysis, by allowing other economically relevant variables to have an indirect price impact.

[153] In the original model Kyle assumes one event, one auction and one informed trader. This allows for a fully developed strategy. However, when there are competitors, such as other informed traders, or a continuous trading scheme, Kyle's equilibrium can no longer be considered valid. Subsequent studies extend Kyle's framework assuming either multiple informed traders (Foster and Viswanathan, 1993; Holden and Subrahmayam, 1992) or continuous auction (Baruck, 1992).

[154] Further studies examine the learning process of traders, known in the REE literature as "herding". Static Bayesian approaches like Kyle (1989), Veronesi (2000) and Chamley (2007) postulate that when information is costly to acquire there is no possible full equilibrium informative price. Banerjee (1992), Bikhchandani et al. (1992) and Chamley (2004) criticize the Bayesian learning and propose the social learning methodology, where fully rational market participants might "herd" towards the wrong direction, following other traders, disregarding existing valuable information. This partially explains financial market crashes, since market failure could be a natural outcome of excessive "wrong herding". In parallel, Jun and Vives (2004) and Vives (2008) examine further the speed of learning, emphasizing on different types of traders and different types of information..

traders, in a way that allows discretionary uninformed traders to learn by observing past order flow. This means that informative trades can be revealed by their characteristics such as transaction size and time of occurrence. Third, information models attribute a double role to past transactions. According to inventory models, they act as indicators of past and future liquidity and therefore they are expected to have only a transitory price impact. In information models, however, they are considered to convey information, which affects prices in the long-term. Finally, according to sequential models, dealers, who can be seen as risk-averse uninformed traders wanting to learn, observe past order flow and they form expectations in order to quote regret free prices.

However, the major disadvantage of inventory and information models is their resilience to game theory. The equilibrium existence depends heavily on the knowledge or assumption, of the game played, such as market structure and market participants. Therefore, unlike these theoretical approaches, empirical market microstructure models have been developed in parallel. An excellent starting point is the seminal study of Roll (1984). He introduces a model, where half spreads can be estimated, assuming martingale behaviour of covariance stationary prices, known as the “covariance” model. Transaction prices are assumed to equal a random walk component, associated with information, and a cost component. This model recognizes only order-processing cost, but it is revolutionary, in the sense that it allows spreads to be estimated using only transaction prices and not using order submissions. Choi et al. (1988), Stoll (1989) and George et al. (1991) extend Roll’s (1984) model to include all three components. Furthermore, after a transformation of the random walk component, Roll’s model can be written as a VAR process. This approach has been followed by Hasbrouck (1991a, 1991b, 1996) and Dufour and Engle (2000b), among others. Another extension of Roll’s model, proposed by Glosten and Harris (1998), introduces another source of uncertainty in the information component - the trade direction, recognizing the informational content of a trade. Further, “trade indicator” models have been developed by De Jong et al. (1996), Madhavan et al. (1997) and Huang and Stoll (1997).

Several empirical aspects have emerged, alongside these models. One of the most important is the effect of trading intensity on prices and spreads. Two measures of trading intensity have been intensively examined, raising both informational and liquidity considerations, associated with past transactions. The first is the transaction size. De Jong et al. (1995) show that trading spreads are a decreasing function of volume. In a similar way, Huang and Stoll (1997) and Ahn et al. (2002) report that order-processing costs decrease with volume. Slightly variant is the study of Stoll (1978), who argues that the relation between transaction costs and volume follows a U-shaped curve. In addition, other studies like Easley and O’Hara (1987), Hausman et al. (1992) and Easley et al. (1997) connect transaction size with information and the type of traders. Along the same lines, Easley and O’Hara (1992) argue that larger volumes are associated with increased presence of information. Furthermore, Chan and Lakonishok (1995) show that after large block trades the permanent, informational price impact is higher. More recently, Angelidis and Benos (2009), extending Madhavan et al.’s (1997) model, re-examine the informational and liquidity impact of transaction size. They report that the information cost component increases with transaction size, while the order-processing and inventory components decrease due to economies of scale.

The next determinant of liquidity refers to time. Even since Glosten and Mirlgrom (1985), Kyle (1985) there is an underlying assumption of the significance of time.[155] However, time itself has not been explicitly incorporated into the microstructure literature until Diamond and Verrecchia (1987) and Easley and O’Hara (1992).[156] From that point time has been used to proxy liquidity or the intensity of trading (see, inter alia, Jasiak and Ghysels, 1998; Engle and Russel, 1997, 1998; Engle, 2000; Renault and Werker, 2002; Manganelli, 2005; Spierdijk, 2004). Furthermore, Engle and Russell (1998), emphasizing the fact that time is irregularly spaced, propose the Autoregressive Duration Model (ACD). Several studies employ standardized durations to account for time variant effects on prices and spreads.[157] Dufour and Engle (2000b), employing a VAR model, examine extensively the role of time on price formation and report that higher trading frequency (they refer to it as higher trading intensity) results in less informative prices, due to increased levels of informed traders (Easley and O’Hara, 1992), and therefore in decreased liquidity.[158] Recently Grammig et al. (2007), trying to

[155] Back and Baruch (2004), show that Kyle (1985) and Glosten and Milgrom (1985) can meet under specific time assumptions. In more details, in a continuous time context, the sequential framework meets the auction setting, with information being resolved in different speeds.

[156] Diamond and Verrechia (1987), in a sequential framework, conjecture that on good news informed traders will always buy more, while upon the arrival of bad news, especially when short selling is not allowed, they cannot short sell. Therefore, long durations are associated with bad news. On the contrary, Easley and O’Hara (1992) maintain that informed traders will always transact upon the arrival of information. Therefore, longer durations are associated with the absence of new price-relevant information.

[157] ACD models model time employing the ratio of actual over expected duration. The so-called standardized durations have been employed by several studies (e.g., Grammig et al., 2007; Ben Sita, 2010) to account for time deformation.

[158] This is on opposition to Kyle (1989), who argues that prices are less informative in cases of decreased competition, because informed traders cannot trade intensively. Easley and O’Hara (1992) postulate that increased market activity, which, consequently, indicates a more competitive market environment, is an indication of informed traders. According to Kyle (1989) this is when prices should be more informative.

provide a connecting link between the structural model of Madhavan et al. (1997) and the VAR approach of Dufour and Engle (2000b), oppose Easley and O’Hara (1992) by arguing that “no trade means no information”. They connect time with trading strategies of uninformed traders and not with information processing as in Dufour and Engle (2000b). They also argue that trades after inactive stages of the market are more informative and therefore they have a larger price impact. More recently, Ben Sita (2010) attributes a dual, informational-permanent and liquidity-transitory, context, similar to Angelidis and Benos (2009), to time and raises the importance of time in the transitory component of the spread.

Two other very important issues, related to trading practice, have been discussed extensively in the literature. First, as Limit Order Markets (LOMs) have been developed, many studies discuss the informational efficiency of these markets (e.g., Brown and Zhang, 1997), the relation between type of orders and traders or spreads (e.g., O’Hara and Oldfield, 1986; Chakravaty and Holden, 1995) or more importantly the trader’s choice concerning the type of trades (see, inter alia, Cohen et al., 1981; Angel, 1994; Kumar and Seppi, 1994; Harris and Hasbrouck, 1996; Harris, 1998; Parlour, 1998; Foucault, 1999; Goettler et al., 2005; Wald and Horrigan, 2005). More recently, Bouchaud et al. (2004, 2006) and Wyart et al. (2008) derive the lagged impact function from the Madhavan et al. (1997) model proposing that the main determinant of the spread is the adverse-selection component while the transitory component causes most of the volatility. Their derivation helps to determine the marginal profit of a limit or a market order. However, the majority of studies seems to agree with the fact that wider (narrower) spreads are associated with limit (market) orders. Second, several studies associate the direction of trade with its size and informational content. Aitken and Frino (1996) show that Buy orders are associated with larger transactions costs, mainly due to information reasons. Likewise, Hedvall et al. (1997) connect the transaction size with information and argue that there are more often information reasons behind large Buys than large Sells.

Focusing on the Carbon market, the studies examining the intraday price formation and the composition of the Bid-Ask spread are limited. Benz and Hengelbrock (2008), estimating Madhavan et al.’s (1997) model, using data from the early stages of the market for EUA futures contracts on the two largest European exchanges, namely ECX and Nord Pool, report a large and significant adverse-selection component along with a small and rather insignificant order-processing component. They also observe a decline in the estimated spreads over the years. In parallel, in an up-to-date study, Frino (2010) reports an increase in liquidity throughout the years and a dramatic increase of liquidity costs. Also, he reports that most transactions have a large information component and concludes that they must have been made by informed traders. Both emphasize the necessity of appropriate regulations in order for the information to be efficiently disseminated and to have operationally large spreads. More recently, Viswanathan (2010) confirms the importance of appropriate regulation and argues that the proposed legislation should serve the ultimate purpose of the market, which is Carbon Emissions management. For that purpose markets should be sufficiently liquid, but at the same time they should allow for market innovation.

Considering the afore-mentioned issues, this chapter employs the empirical findings discussed in the previous two chapters. Market makers and limit order traders, who are seen as market makers in one side of the spread, are recognised to formulate expectations about the price impact of future transactions and they quote their prices accordingly. Using the STM-ACD model, they can forecast expected duration and they can attribute a probability of the next transaction to be informed. They can do that by forecasting the level of activity the market is expected to be in. This way they can determine how long are they expected to be exposed to information related risk, and what the price impact of the incoming trade is expected to be.

A key aspect of this process is the dual pricing impact of trading intensity, which unlike previous studies is modelled simultaneously, by allowing expectations to affect the information and liquidity component of price change. This dual impact can be further interpreted as a natural measure of market sensitivity towards information and liquidity. These sensitivities are expected to vary over time and across different market participants. Consequently, their price impact is expected to depend on current market conditions. This could explain the conflicting results in the literature, concerning the price impact of transaction time and size.

In more detail, a new dynamic expectations structural model for intraday price changes is proposed, extending the existing literature in the following ways. First, it deviates from a static approach of modelling intraday prices and Bid-Ask spread, in the sense that it allows price components to vary according to trading activity. The information and liquidity components are revised after trading intensity fluctuations. Second, price revisions depend on the, Bayesian, learning process of market participants, who condition their quoted prices upon expectations about the price impact of future transactions. Consequently, price and spread components are modelled as continuous latent variables that depend on past trading history, public information and market participants' learning ability. Third, the dynamic character of the model provides a framework to measure the information and liquidity pricing impact of a trade, in a way that the profitability of a market or limit order strategy can be maximized. These components are measurable and they are revised after every transaction. Investors can formulate expectations and then take an appropriate market position. Finally, this model investigates further the intraday price formation of Carbon allowances, recognising the pricing impact of market participants' behaviour. This can have numerous implications to regulation and trading in EU ETS.

In more detail, the present study extends Madhavan et al.'s (1997) model towards the directions of Grammig et al. (2007), Angelidis and Benos (2009) and Ben Sita (2010), proposing a new dynamic expectations structural model for intraday returns.[159] The analytical focus is upon dealers' behaviour in the Carbon market. According to the literature, dealers are recognized as uninformed market participants, who continuously observe order flow in order to extract information signals. They use these signals to formulate expectations for future transactions. The quoted spreads are considered to be regret free, depending on these expectations. A key element of this analysis is the

[159] The choice of this particular model seems to be the obvious choice for the empirical examination of the issues addressed. First, it is a structural model, which is more appropriate in order to study the data generation process of prices, under the analytical spectrum of how market participants act. The restriction, though, is that it is a one period model, since only the last transaction directly affects the price formation. However, since investors' acts induce serial dependence on the price residuals, an agnostic VAR, with non-zero serial moments of order greater than two, would be mis-specified. Although VAR is a great analytical tool, especially when long memories, time impulse functions and forecasting are of interest, in a structural specification, disturbances should possess complicated and non-intuitive properties (Greene, 2002). Second, compared to Huang and Stoll's (1997) structural model, which shares a very strong resemblance, the MRR model appears to have gained considerable popularity. Huang and Stoll (1997) propose a "complete" model, in the sense that it includes covariance and trade indicator models as special cases and that it decomposes the spread into all three widely recognized components. However, several studies, like Bollen et al. (2004), Kumar (2004) and Angelidis and Benos (2009) criticize the large order-processing component of the spread. Easley et al. (1996) are in line with Huang and Stoll, maintaining that information costs are likely to be low in a dealer-dominated market, since informed traders are less likely to trade with dealers. However, the studies of Benz and Hengelbrock (2008) and Frino (2010) report rather small and sometimes insignificant order-processing costs. In addition, Huang and Stoll's (1997) model, derives the cost components indirectly after estimating half spread and covariances first. Therefore, MRR model seems to be more appropriate for this analysis, because although it sacrifices the generality of Huang and Stoll's (1997) or the longer memory of the VAR models, it allows for a structural approach that estimates the various spread components in one step. In addition, Madhavan et al.'s (1997) model shares a very strong resemblance with Glosten and Harris's (1988) (GH) model. However, in the GH model trade sign and size are considered to be exogenous to the trading process, and therefore strong distributional assumptions need to be made. On the contrary, with the MRR model and the extension proposed in the present analysis, trading intensity, in terms of both trade size and time, and trade sign are derived from the trading process.

intensity of trading and its ability to identify the presence of informed traders, which has been analyzed in the previous chapter. Trading intensity is employed to decompose the spread into its three, widely recognized components, as a continuously updated variable. Therefore, spread and its components can vary across transactions. This helps in identifying a unique spread for every transaction, or at least for every similar event, as it is defined by different levels of trading intensity. Trading intensity is also allowed to have an impact on both the adverse-selection and liquidity components. In addition, following the early literature of inventory models, the transitory, liquidity component is enhanced with two other risk aversion measures. This "fourth" component accounts for the expected number of informed traders and the expected price volatility, taking into account the time dealers are expecting to be exposed to these types of risk.

The proposed model conjectures that trading intensity can explain various issues related to the trading process and the empirical findings support its adequacy. One of the main findings of this study is that the spread and spread components have indeed an intraday pattern, similar to the findings in Madhavan et al. (1997). According to the empirical analysis in this chapter, these intraday patterns can be attributed to trading intensity variations.[160] Trading intensity seems to be positively related to the information component. In particular, when the next transaction is expected to be larger or faster, this can be interpreted as higher informational content, dealers appear to increase the spread, which is consistent with Dufour and Engle (2000b). In contrast, when trading intensity is taken into account, the order-processing cost appears to be considerably larger with a negative relation to trading intensity. A potential explanation for this is either the economies of scale, similar to Angelidis and Benos (2009), or the illiquid character of the market. Also, volatility seems to be the prevailing risk component increasing the spread (Bollen et al., 2004).[161]

Another main finding is the confirmation of the positive relation between limit orders and spread width. However, it appears that a limit order strategy could be profitable

[160] Several studies have examined the intraday, seasonal patterns of either quoted or estimated spreads (see, inter alia, Kim and Ogden, 1996; Madhavan et al., 1997; Huand and Stoll, 1997; Brockman and Chung, 1998; Chan, 2000; Ahna et al., 2002; Gwilym and Thomas, 2002; Silva and Chavez, 2002; Pinder, 2003; Ke et al., 2004). The results have been conflicting and they seem to be relevant to microsturctural aspects of each market. The present analysis provides an insight into these variations, conjecturing that they might depend on variations in trading patterns that can be adequately summarized by trading intensity.

[161] These findings confirm and explain various issues in the literature. First, the relation between volatility and spreads proposed by inventory models seems empirically reasonable. Second, the conflicting results about transaction size (Huang and Stoll, 1997; De Jong et al., 1995, 1996; Ahn et al., 2002) and time (Kyle, 1989; Engle, 2000) and their relation to spreads can be explained by the different attributes of trading intensity on the information and liquidity components.

only when trading intensity is low. Otherwise, the information component indicates that the midpoint deviation would be larger than the spread and thus a limit order should be preferable. Finally, Buys and especially the large ones are associated with a higher adverse-selection component and wider spreads.

These findings are particularly relevant to various aspects of trading process in EU ETS. First, trading practices can be improved, since investors could measure the price impact of their trades, which might vary over time, more precisely. They can account for behavioural aspects of trading in real time and they can formulate trading strategies that take into account the actions of other investors. These actions can be included in the expectations' equation and thus can be explicitly modelled. This provides a promising framework that connects the theoretical approach of "inventory" and "information" models with the empirical flexibility of "trade indicator" models and the forecasting ability of "VAR" models.

Further, this practice can also be beneficial to investors that posses price unresolved information, since they can adjust their trading to current liquidity levels and sensitivity towards new information, and maximize their profits. These traders can measure the post-trade price impact of their actions and they can adjust their strategies accordingly in order to minimize their visibility and maximize their profit.

Second, market making can be further improved, since this model can indicate when market orders can be profitable. Dealers can adjust spread width according to liquidity and information and they can become more efficient in managing both. This would result in narrower spreads and further contribute to market development and maturity.

Finally, a better understanding of trading practices and their pricing impact can improve regulation and monitoring, and thus the market can be more efficient in achieving its goal, to reduce emissions. This model proposes a natural measure of market sensitivity towards information and liquidity, which are both highlighted in the Carbon market literature. Regulatory authorities can develop policies that can manage both and allow the market to find a balance between "market innovation" and liquidity.

This chapter is organized as follows. The next section describes methodological issues, while the one following discusses the empirical findings. A brief summary is provided at the end of the chapter.

6.2 Methodology

6.2.1 Madhavan et al. (1997) (MRR) Model

Madhavan et al. (1997) propose a structural model of the quote and return generating mechanism in a hybrid market (NYSE), which henceforth will be referred to as the MRR model and in which there are two broad categories of trader; liquidity providers (i.e., dealers and investors that submit limit orders) and liquidity takers (i.e., investors that submit market orders).[162] They postulate that liquidity providers, after extracting sufficient information from the market, quote regret free prices at which they are willing to trade.

Let $\tilde{v}_t$ denote the "true" value of a risky asset at some point in time t. This is the full information expected present value of the asset's future cash flows, given the full information set Φ_t. This value can change over time due to variations in expectations (future cash flows or the discount rate) according to variations in public information. Furthermore, let $\mu_t = E[\tilde{v}_t|H_t]$ be the post-trade conditional expectation of $\tilde{v}_t$, with $H_t \subseteq \Phi_t$ being the available information set at time t.[163]

They assume that the change in beliefs, concerning the fundamental value of the asset, depends on changes in the available information set H_t. Contrary to Roll (1984), they follow the example of Glosten and Harris (1998), who introduce the trade sign as another source of uncertainty in determining the available information sets. Trade sign is defined as:

$$q_t = \begin{cases} +1 & when\ the\ transaction\ is\ Buyer\ initiated \\ -1 & when\ the\ transaction\ is\ Seller\ initiated. \end{cases} \tag{6.1}$$

They postulate that price changes are associated first with new public information (ε_t), which is not associated with trading, and second with innovations in order flow. The

[162] They mention that, although they do not explicitly model the limit order book, they can still extract the spreads from transaction prices, by assuming that limit orders submissions are one side market making. This, given the different scale, seems to be fairly similar to the Carbon market, where dealers have been introduced to enhance liquidity in a limit order market.

[163] Hasbrouck (2007) explains that, when data sets of high frequency are employed, the estimated mean return is small relative to the estimation error of its arithmetic mean. Therefore, it is a common practice in the microstructure literature to drop the mean return from the models, implicitly assuming that it is zero. Zero is a biased estimator mean return, but its estimation error is lower than that of the arithmetic mean. This makes μ_t, which is the current "fair" value the best future estimator, $E[\mu_{t+1}|H_t] = \mu_t$. Then the conditional expectation $\mu_t = E[\tilde{v}_t|H_t]$ is a martingale with respect to the available information set H_t. H_t is unique for every market participant and limit order traders' quoted prices depend on it. When H_t consists of all available public information Φ_t, then $\mu_t = \tilde{v}_t$ and is called the "fundamental" value of the asset or, according to the asset pricing literature, the "efficient" price.

timing of the events is assumed to follow this order. After a transaction in $t-1$, the "post-trade efficient" price is μ_{t-1}.[164] Then, public information arrives as the realization of ε_t. At the same time, dealers formulate their expectations about order flow and trade direction, given the previous transaction, $E[q_t|q_{t-1}]$, assuming that Buys and Sells are equally likely ($E[q_t]=0$) and that q_t is autocorrelated of order one ($Cov(q_t, q_{t-k}) = 0\ for\ every\ k \geq 2$).[165] When the next transaction, in t, occurs, it is either Buyer or Seller initiated. The realization of q_t, if it is different from its expectation, measures the size of the innovation in order flow, $(q_t - E[q_t|q_{t-1}])$. The new efficient price will then be μ_t. Therefore, the change in beliefs can be mathematically formulated as:

$$\mu_t - \mu_{t-1} = \theta * (q_t - E[q_t|q_{t-1}]) + \varepsilon_t, \qquad (6.2)$$

where μ_t is the post-trade efficient price, incorporating all available, public (ε_t) and trading information up to time t and $(q_t - E[q_t|q_{t-1}])$ is the term that measures the size of the innovation in order flow. Informed traders expose their information through realized strategies. They buy (sell) when the asset is under-priced (over-priced). Therefore, a trade initiation, different from the expected, denotes information that dealers and other uninformed traders do not possess and which has yet to be incorporated into prices. $\theta \geq 0$ is a parameter that measures the responsiveness of price towards the innovations in order flow. Madhavan et al. (1997) argue that this is a natural measure of information asymmetry and therefore it accounts for the permanent, information price impact of trades.

Moreover, given that dealers need to quote prices, at which they are willing to trade, prior to the event at t, and that they are risk averse, they need to incorporate the price impact of the next transaction at t into their quotes. Therefore they quote two prices: the price they are willing to buy at, the Bid price, and, another at which they are willing to sell, the Ask price. In this way, they ensure that they quote regret free prices according to all available information, namely public information and trading information derived from the last trade. Denoting the transaction price at time t by p_t, the Bid and Ask can be formulated as:

---

[164] Please note that the index *t* stands for a particular event, such as transaction, and not fixed calendar time. Hasbrouck (2007) argues that, when detailed data sets with precise time stamps are employed, it is more accurate to mark events by the time of their occurrence and not by calendar time. This allows for a deeper microstructure analysis. The problem that arises, though, in this case is the modelling of duration, defined as the irregularly spaced time between two consecutive events.

[165] Madhavan et al. (1997) follow Glosten and Milgrom (1985), who argue that the revision in beliefs is directly proportional to the actual order flow, assuming that it is autocorrelated.

$$p_t^a = \mu_{t-1} + \theta(1 - E[q_t|q_{t-1}]) + \varphi + \varepsilon_t, \quad (6.3)$$

$$p_t^b = \mu_{t-1} - \theta(1 + E[q_t|q_{t-1}]) - \varphi + \varepsilon_t. \quad (6.4)$$

where $\varphi \geq 0$ is a parameter that measures dealers' per unit cost of processing the transaction and for supplying liquidity, which is assumed to be constant over time and for all levels of liquidity. Then, assuming that the transaction occurs either on the Ask or on the Bid, the transaction price can be written as:

$$p_t = \mu_t + \varphi q_t + \xi_t, \quad (6.5)$$

where ξ_t captures the rounding error, induced by price discreteness.[166] Combining Eq. (6.2) with (6.5), the theoretical model is given by:

$$p_t = \mu_{t-1} + \theta(q_t - E[q_t|q_{t-1}]) + \varphi q_t + \varepsilon_t + \xi_t. \quad (6.6)$$

However, since μ_{t-1} is latent and $E[q_t|q_{t-1}]$ cannot be observed, some transformation needs to be applied in order for the model to be estimated. Madhavan et al. (1997) assume that order flow is first order autocorrelated. They denote by ρ the first order autocorrelation, which is given by $\rho = \frac{Cov[q_{t,}q_{t-1}]}{Var[q_{t-1,}]}$. Consequently, $E[q_t|q_{t-1}] = \rho * q_{t-1}$. In addition, they substitute μ_{t-1} using Eq. (6.5) and this gives the testable equation:

$$\Delta p_t = (\theta + \varphi)q_t - (\rho\theta + \varphi)q_{t-1} + \big(\varepsilon_t + (\xi_t - \xi_{t-1})\big). \quad (6.7)$$

This model decomposes price changes into an information, and thus permanent, component, which is measured by θ and is related to the informational content of order flow; a transitory component, which is measured by φ and is attributed to trading frictions; and a noise component induced either by public information (ε_t) or price discreteness (ξ_t). Further, after estimating the parameter vector $\beta = (\theta, \varphi, \rho)$, they draw inferences about the implied spread ($p_t^a - p_t^b$):

$$S^{Implied} = 2(\varphi + \theta), \quad (6.8)$$

the effective spread (cost of a round trip):

[166]According to Hasbrouck (2007), Madhavan's (1997) model and trade indicator models in general constitute a generalized form of Roll's (1984) model. In covariance-based models, the efficient price is seen as a random walk, where here a trade related component is added. This general formulation is quite common in the empirical microstructure literature. In addition, in the Carbon market the minimum tick is €0.01. Therefore ξ_t and its variance, σ_ξ^2, are expected to be relatively small.

$$S^{Effective} = 2\varphi + \theta, \tag{6.9}$$

and the return variance:[167]

$$Var(\Delta p_t) = \underbrace{\sigma_\varepsilon^2 + 2\sigma_\xi^2}_{\substack{Public\ Information \\ and\ Discretness}} + \underbrace{(1-\rho^2)\theta^2}_{\substack{Asymmetric \\ Information}} + \underbrace{2(1-\rho)\varphi^2}_{\substack{Transaction \\ Cost}} + \underbrace{2\varphi\theta(1-\rho^2)}_{Interaction}. \tag{6.10}$$

Both, estimated, spread measures include a permanent and a transitory component, associated with information and liquidity, respectively. Variance is decomposed into a Public Information, a Price Discreetness, an Asymmetric Information, a Trading Cost and an Interaction between Asymmetric Information and Trading Cost component.

6.2.2 A Dynamic Expectations Structural Model

Foundations of the model

Although the MRR model is a very simple and convenient way to estimate the informational and liquidity components of the spreads and has gained popularity, it still has some major drawbacks. First, it is "incomplete", in the sense that it summarizes order-processing and liquidity cost components in one parameter. Second, as Grammig et al. (2007) postulate, although Madhavan et al. (1997) recognize the spread as a random variable and that φ and θ exhibit intraday patterns, their model does not allow for intraday variations. In order to overcome this problem in the MRR model, they estimate the model various times for different trading periods during the day. However, this violates the usual assumption of continuous trading, when UHF data is employed. Third, it only provides estimations of the spread for unit (transaction) size. Angelidis and Benos (2009) provide evidence that this approach is "restrictive", at least in an emerging market such as the Athens Stock Exchange. They argue that the information component should be directly proportional to the traded volume, while the transitory component should include a liquidity related variable. They provide evidence that the information component increases with volume, which is consistent with Dufour and Engle (2000b), while the transitory component decreases as transaction size increases.

[167] The exact formulas for the implied and the effective spread can be derived by subtracting Eq. (6.4) from (6.3). On the contrary, for the variance Eq. (6.6) is used under the assumptions that order flow is autocorrelated of order one and that Buys and Sells are equally likely, while both innovations, due to public information and price discreteness, are serially uncorrelated. For more information and the exact specifications please refer to the original paper.

They attribute this variation to economies of scale. Fourth, Grammig et al. (2007) and Ben Sita (2010) argue that time plays an important role in price formation and should not be ignored. They both argue that duration conveys information and therefore should permanently revise prices, while at the same time it is an indication of liquidity that should have a transitory price impact.

The present analysis extends the MRR model in various directions of effects that have been discussed in the literature but that have not been implemented in a dynamic structural context. The role of trading intensity is examined extensively. In accordance with the previous chapters, it is measured as transaction size per unit of time, summarizing the effects of transaction size and time. Following the previous chapter, its different regimes and their connections to different types of traders are of particular importance. Moreover, trading intensity is recognized as another price-relevant factor, along with the trade initiation variable, that dealers need to manage explicitly and formulate expectations for. According to the literature, trading intensity is allowed to have an impact on revisions in beliefs, which is permanent, as well as a transitory on the liquidity component. The source of this impact is assumed to be the formulation of expectations concerning future levels of trading intensity and more importantly the potential regime that it might be in. Consequently, every transaction affects the expected trading intensity and therefore the quoted prices and spreads. This way, φ and θ are modelled as functions of trade-related variables and revise their values after every trade. Therefore, contrary to the MRR model, a single estimation is sufficient for calculating the spread after every transaction or over different trading periods. In addition, dealers' risk aversion is taken into account, and, together with trading intensity, is used to decompose spreads into all three recognized components. Finally, the dynamic character of the model indicates different regimes of trading intensity where different types of order (namely market or limit) are preferred in terms of profitability.

More precisely, price revisions are assumed to be driven by both a permanent and a transitory component. The permanent component is related to information and changes in expectations about the value of the asset. These changes are attributed partially to exogenous non-trading-related public information shocks, and partially to the trading process. The innovation in order flow is assumed to reveal private information that has not been incorporated into prices yet. In the model presented in this chapter this effect is variant, depending on transaction size and time of trade. It is allowed to vary according to what dealers expect concerning the future state of the market, in terms of liquidity, informed trading and price volatility. The same variables are also allowed to have an impact on the transitory component, which represents the dealers' cost for supplying liquidity. Consequently, after every trade, dealers revise their expectations about the "fair" value of the asset, this is the permanent price component, and the future liquidity, which is the transitory price component. This is reflected in their quoted prices.

The model

Similar to the MRR model, the revision in beliefs concerning the post-trade efficient value of the asset depends on public information shocks and innovation in the trading process. The main source of uncertainty is the surprise in order flow. However, in this analysis the sensitivity of price to the size of the innovation is assumed not to be constant. The characteristics of the trade, as well as the contemporary state of the market, are assumed to be the driving forces of the sensitivity variations. The size of the transaction and the time of its occurrence are used as the characteristics of the transaction, while market conditions are summarized by the current level of volatility and the number of informed traders.

Dealers, after observing the previous transaction at $t-1$, revise their beliefs about the fair value of the asset μ_{t-1}. Then, they need to quote a Bid and an Ask price for the next transaction. They know, though, that the post-trade fair value of the asset will be affected by the next transaction. Therefore, in order to be able to set regret-free prices, they need to take into account any post-trade effects. However, they cannot be sure of the size of the next incoming trade, or of when it will happen, or of the contemporaneous level of volatility or private information. They can only speculate.[168]

This is the basic assumption of the proposed model. Dealers still condition their quotes on the sign of the incoming trades, but this is assumed constant only for a pre-specified depth, in terms of liquidity, as it is summarized by transaction size and time. They are

[168] Previous studies, such as De Jong et al. (1995,1996), Huang and Stoll (1997) and Angelidis and Benos (2009), that have examined the price impact of trade size, have used the actual values (i.e., size of the limit order) of the variables and not the expectations (i.e., size of the incoming transaction). The main reason for the difference with this analysis is the market examined. They examine a pure Limit Order Market (LOM), where there are not designated market makers and the spread is determined by submitted limit orders, compared to a hybrid market such as the Carbon market. These models emphasize the midpoint of the best Bid and Ask prices and its intraday formation. In a LOM, the size and the time of the submission of the order are known, at least to the investor, and they can be used directly. On the contrary, in a hybrid market, emphasis is upon the transaction prices and the price setting of market makers. Although they can observe the submission of limit orders, which are assumed to be market making in one side of the spread, they cannot be sure of the size or the timing of an incoming market order. Therefore, they can only formulate speculations based on information currently available.

assumed to formulate expectations concerning the size of the next transaction, the time of its occurrence and the future levels of price and information uncertainty, in a continuous manner, updating themselves continuously. The results of the analysis in the previous chapter show that trading intensity carries information concerning informed trading. Associating this with trading intensity expectations, dealers can, indirectly, speculate about the probability of informed trading, which definitely has a price impact. More specifically, dealers formulate expectations concerning the future levels of the intensity of trading. Their expectations might lie in different regimes, which might indicate a different type, informed or uninformed, of incoming trade. Then, they quote their prices accordingly. Furthermore, trading intensity is assumed to play a dual role in intraday price formation. Apart from the informational part, it is a natural measure of the liquidity of the market. This concept is of great importance in the inventory models and it has been reported extensively in the literature as a main determinant of the transitory component of the spread. Consequently, in this study, different regimes of trading intensity are allowed to revise the formation of dealers' transitory cost, as well as their beliefs. Finally, a potentially permanent and/or transitory, or both, effect of risk is also examined.

Following the analysis in Chapter 5, let $s_t = \frac{v_t}{d_t}$ denote trading intensity, where v_t is transaction size and d_t is raw duration.[169] Then, $E[s_t|H_{t-1}]$ is the expected value of s at time t and H_{t-1} consists of all available information up to time $t-1$. s_t is modelled as a martingale process with respect to H_{t-1}. It is conjectured to be affected by random shocks ($e_t, assumed\ to\ be\ independent\ of$ ε_t), its previous values ($\sum_{j=1}^{J} s_{t-j}$), previous returns ($\sum_{i=1}^{I} r_{t-i}$), previous trade directions ($\sum_{k=1}^{K} q_{t-k}$) and other stylized facts, such as whether the transaction is informed ($\sum_{l=1}^{L} I_{t-l}$) or OTC ($\sum_{m=1}^{M} D_{t-m}$).[170] D_t and I_t are dummies with the following values: [171]

$$I_t = \begin{cases} +1 & \text{when } s_t > s_2 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad D_t = \begin{cases} +1 & \text{if transaction is OTC} \\ 0 & \text{otherwise.} \end{cases} \tag{6.11}$$

Then, $E[s_t|H_{t-1}]$ and can be formulated as:

[169] In order to simplify the estimation process, the diurnally adjusted trading intensity is used. This way no explicit modelling of seasonalities is needed.

[170] s_t is modelled as an Autoregressive process in order to account for time dependence, while other variables have been added to indirectly account for the price impact of previous trades. More specifically, an abnormal return or order initiation might trigger the trading process. On a similar fashion, when an OTC holder or an informed trader enters the market, the intensity of trading might change accordingly.

[171] s_2 is the threshold value, unique for each data set, that distinguishes informed versus uninformed trades. The value is taken from the STM-ACD model of the previous chapter.

$$E[s_t|H_{t-1}] = s_t = f(s,r,q,I,D) = \mathrm{c}_1 + \sum_{j=1}^{J} c_{2,j}s_{t-j} + \sum_{k=1}^{I} c_{3,k}r_{t-i} \\ + \sum_{i=1}^{K} c_{4,i}q_{t-k} + \sum_{l=1}^{L} c_{5,l}I_{t-l} + \sum_{m=1}^{M} c_{6,m}D_{t-m} + \mathrm{e_t}, \tag{6.12}$$

where $c = [c_{h,q}],\ h = 1,\ldots,6\ and\ q = j,\ldots,m,$ is a vector of parameters to be estimated. For consistency with the foundations of the model we assume first order autocorrelation throughout. This is $j = k = i = l = m = 1.$ [172]

Furthermore, dealers are assumed to formulate expectations about the future level of risk. Two sources of risk are recognized, price volatility and presence of informed traders. Both are measured as moving averages of the last fifteen minutes.[173] This is:

$$E\left[\sigma_{p,t}^2|H_{t-1}\right] = \sum_{n=1}^{N}\left(p_{t-n} - \sum_{n=1}^{N} p_{t-n}\Big/N\right)^2\Big/N \tag{6.13}$$

$$E[\tilde{I}_t|H_{t-1}] = \sum_{n=1}^{N} I_{t-n}\Big/N \tag{6.14}$$

where $E\left[\sigma_{p,t}^2|H_{t-1}\right]$ is the expected price volatility, and $E[\tilde{I}_t|H_{t-1}]$ is the expected number of informed traders at time t given all information available at $t-1$.

Given the above, the revision in post-trade beliefs could be formulated as in (6.2) but with time varying θ, θ_t:

$$\mu_t - \mu_{t-1} = \theta_t(q_t - E[q_t|q_{t-1}]) + \varepsilon_t, \tag{6.15}$$

[172] A basic assumption of the original model is that a return's covariances of second order and higher are essentially zero. This indicates a short memory of order one. The analysis presented here, however, provides a basis for incorporating older information indirectly through formulating expectations, into price formation.

[173] This assumption has been made for simplicity. Both expectations could be modelled as latent processes with other exogenous factors and public innovations determining the conditional mean, similar to trading intensity. However, that would require a quite complicated modelling of the covariances of the innovations, or a quite strong assumption that there is no inter-relation. Another approach could be the use of a Heterogenous Auto Regressive Realized Volatility (HAR-RV) (see, inter alia, Corsi et al., 2008; Chevallier and Sévi, 2010) framework for both risk measures. However, that would require the assumption that both are exogenously determined. Although it could be reasonable for informed trading, it would be a very strong assumption for price volatility. In contrast, the moving average framework is a relatively simply approach that does not require any exogeneity assumption.

where

$$\theta_t = \theta\left(s, \tilde{I}, \sigma_p^2, \psi\right) = \theta_1 + \sum_{\pi}^{3}\left(\theta_{2,\pi} * I_{\pi,t}\right) E[s_t|H_{t-1}] + \left\{\theta_3 E[\tilde{I}_t|H_{t-1}] + \theta_4 E\left[\sigma_{p,t}^2|H_{t-1}\right]\right\}\psi_t, \quad (6.16)$$

where $\pi = uninformed(1), fundamental(2)\ and\ informed(3)$, θ's are parameters to be estimated, ψ_t is the expected duration given by an STM-ACD model and $I_{\pi,t}$ is a variable that indicates the regime that the expected trading intensity is in at time t.[174] $I_{\pi,t}$ can be defined as:

$$I_{\pi,t} = \begin{cases} I_{uninformed} = 1 & when\ E[s_t|H_{t-1}] < s_1 \\ & and\ 0\ otherwise \\ I_{fundamental} = 1 & when\ s_1 < E[s_t|H_{t-1}] < s_2 \\ & and\ 0\ otherwise \\ I_{informed} = 1 & when\ E[s_t|H_{t-1}] > s_2 \\ & and\ 0\ otherwise. \end{cases} \quad (6.17)$$

This formulation is similar to the MRR model, but θ is now decomposed in trading intensity and risk effects. In particular, the coefficients $\theta_{2,\pi}$ measure price responsiveness to innovations in order flow, linked to dealers' expectations of trading intensity. According to Dufour and Engle (2000b), higher trading intensity leads to higher price revisions, due to higher informational content of trades. This is also consistent with the findings of the previous chapter, where higher trading intensity was found to be associated with informed trading. Consequently, when dealers expect higher trading intensity, they should expect higher price revisions for a given surprise in order flow. Therefore, the sum of θ_1, which statistically is the long-term trend of price sensitivity, and θ_2, which is the weighting of trading intensity, should be higher and more positive in the informed regime, compared to the other two regimes. In addition, besides the magnitude, the sign of the coefficient (θ_2) is important as well. A positive

[174] Unlike the risk measures employed, following Engle and Russell (1998) and Dufour and Engle (2000b) duration is assumed to be exogenous to price formation and even though this is a strong assumption it might find justification with data set employed. The market is relatively new and there is increasing trend on the number and size of transactions. In addition, the assets examined are Futures contracts and they are expected to exhibit seasonal patterns. These patterns are exogenous to pricing. In this study duration is assumed to be exogenous for simplicity reasons. However, the literature (e.g., Grammig and Wellner, 2002; Grammig et al., 2007) proposes simultaneous modelling of durations and returns, if there is strong evidence of endogeneity.

sign (expected) would indicate an increased price effect of larger or faster transactions and would be consistent with the propositions of Easley and O'Hara (1992) and Dufour and Engle (2000b).[175] A negative sign would indicate that liquidity is more important than information, since price variations are larger when the market is less liquid.

The last term, third, in Eq. (6.16) is a dual measure of risk. First, dealers are assumed to observe the market, update themselves continuously, and interpret information signals according to the analysis in the previous chapter. Therefore, they can identify past informed transactions, or otherwise the presence of informed traders in the market. This allows them to manage their pricing according to how risky they believe the market is due to the presence of these traders. In addition, expected price volatility is another important risk factor that dealers need to manage. Higher volatility means that price variations are large. A significant part of these variations should be explained by information asymmetry, and, therefore, the price impact of risk should not be ignored. Both risk factors, informed trading and price volatility, are multiplied with expected duration. This term captures how risky the market is expected to be and for how long dealers can expect to be exposed to these risk factors. The sum of θ_1 and θ_3, and/or θ_4, measures the responsiveness of prices to changing market conditions, as they are measured by risk and order flow innovations. Positive signs, which are expected, would indicate higher price variations for a given surprise in order flow.

Along the same lines, the transitory price component φ in Eq. (6.6) can also be decomposed to account for trading intensity and risk variations:

$$\varphi_{\mathrm{t}} = \varphi\left(s\,, \tilde{I}\,, \sigma_p^2, \psi\right) = \varphi_1 + \sum_{\pi}^{3}\left(\varphi_{2,\pi} I_{\pi,t}\right) E[s_t|H_{t-1}] \\ + \left\{\varphi_3 E[\tilde{I}_t|H_{t-1}] + \varphi_4 E\left[\sigma_{p,t}^2|H_{t-1}\right]\right\} * \psi_t, \tag{6.18}$$

where φ's are parameters to be estimated. More specifically, φ_1 measures the long-term average cost component and the constant cost component (φ) in the MRR model. It is constant over time and market conditions and is close to the definition of order-

[175] Easley and O'Hara (1992) maintain that increased volume has a larger price impact due to the increased presence of informed traders. On the same vein, Dufour and Engle (2000b) argue that price variations are larger upon increased trading frequency. A positive θ_2 coefficient would indicate that price variations should be larger for a unit change in order flow innovation when the expected trade size and trading frequency are expected to be higher. This result would empirically support both studies.

processing cost, which is usually fixed. In addition, the coefficients $\varphi_{2,\pi}$ measure the transitory price impact of trading intensity.

According to Benos and Angelidis (2009), a negative sign should be expected for the parameters $\varphi_{2,\pi}$, due to economies of scale. Since trading costs are mainly fixed, increased trading volume should decrease the per unit cost. In very liquid, and rather competitive markets this component should equal the cost of trade realization. Considering that the Carbon market is rather illiquid, larger size and faster transactions are expected to improve dealers' inventory positions and, hence, they should contribute to lower trading costs. Therefore, these parameters account for the liquidity component of the spread. This effect opposes the permanent, informational impact of trading intensity and the spread is determined by the relative magnitude of $\theta_{2,\pi}$ and $\varphi_{2,\pi}$. This would provide at least a partial explanation of the conflicting results in previous studies concerning transaction size, time and price.[176] Finally, risk and time of exposure are expected to contribute to increasing spreads. According to the assumption of risk aversion, dealers should require higher compensation when they are exposed to risk over longer periods or to high levels, or both. This accounts for the risk-aversion component of the spread.

Furthermore, following the MRR model, while incorporating the extensions proposed above, dealers quote the following, regret free prices:

$$p_t^a = \mu_{t-1} + \theta_t(1 - E[q_t|q_{t-1}]) + \varphi_{\mathrm{t}} + \varepsilon_t, \tag{6.19}$$

$$p_t^b = \mu_{t-1} - \theta_t(1 + E[q_t|q_{t-1}]) - \varphi_{\mathrm{t}} + \varepsilon_t. \tag{6.20}$$

The general form could be written as:

$$p_t = \mu_{t-1} + \theta_t(q_t - E[q_t|q_{t-1}]) + \varphi_{\mathrm{t}} q_t + \varepsilon_t, \tag{6.21}$$

Finally, after the appropriate transformations and after gathering like terms, the joint model of returns and trading intensity can be written as:

[176] Concerning the impact of volume on prices and spreads, previous studies, such as Stoll (1978), Easley and O'Hara (1987), Hausman et al. (1992), De Jong et al. (1995), Huang and Stoll (1997), Easley et al. (1997) and Ahn et al. (2002), provide conflicting results that appear to depend on specific stylized facts of each market and/or the period examined. In parallel, another stream of literature (e.g., Easley and O'Hara, 1992; Dufour and Engle, 2000b; Grammig et al., 2007; Ben Sita, 2010) discuss the implications of time on intraday price and spread formation, without concluding on a universally accepted principle. The model proposed in this study conjectures that these differences could be explained by the dual effect of trading intensity, which is expected to be different on the information and liquidity components.

$$E[s_t|H_{t-1}] = s_t = f(s, r, q, I, D) \tag{6.22}$$

$$\Delta p_t = (\theta_t + \varphi_t)q_t - (\rho\theta_t + \varphi_{t-1})q_{t-1} + \left(\varepsilon_t + (\xi_t - \xi_{t-1})\right). \tag{6.23}$$

In this formulation, θ_t and φ_t are the values at time *t*, of continuously updated functions of expected trading intensity ($E[s_t|H_{t-1}]$) and expected risk. They measure the price sensitivity towards asymmetric information and trading costs. These sensitivities are revised after every transaction. Therefore, intraday patterns of spread components can be calculated non-parametrically. In addition, when $\theta_{2,\pi} = \theta_3 = \theta_4 = \varphi_{2,\pi} = \varphi_3 = \varphi_4 = 0$, the joint model nests the original MRR model. Furthermore, by restricting to zero the autocorrelation of order flow ($\rho = 0$), the model can be a special case of that of Huang and Stoll (1997). Moreover, for $\theta_{2,\pi} = \theta$, $\varphi_{2,\pi} = \varphi$, $I_\pi = 1$, $E[s_t|H_{t-1}] = s_t$, $d_t = 1$ and $\theta_1 = \theta_3 = \theta_4 = \varphi_3 = \varphi_4 = 0$ the model simplifies to that of Angelidis and Benos (2009).

Following the MRR model, two measures of spread can be derived from the above specifications, subtracting Eq. (6.20) from (6.19).

The first is the implied spread ($p_t^a - p_t^b$):

$$S_t^{Implied} = 2(\varphi_\mathrm{t} + \theta_t). \tag{6.24}$$

The second is the effective spread (cost of a round trip):

$$S_t^{Effective} = 2\varphi_\mathrm{t} + \theta_t. \tag{6.25}$$

The implied spread could further be decomposed into the following components (the formulas give the proportion of each component, relative to implied spread):

Adverse-selection (Permanent) Part	Trading Frictions (Transitory) Part	
$\theta_t/2(\varphi_\mathrm{t} + \theta_t)$	$\varphi_\mathrm{t}/2(\varphi_\mathrm{t} + \theta_t)$.	(6.26)

Considering Eqs. (6.16) and (6.18), the Trading Frictions Part can be further decomposed to the order-processing cost component, which can be written as $\varphi_1/2(\varphi_\mathrm{t} + \theta_t)$, the liquidity component $\left\{\sum_\pi^3 \left(\varphi_{2,\pi} * I_{\pi,t}\right) E[s_t|H_{t-1}]\right\}/2(\varphi_\mathrm{t} + \theta_t)$ and the risk-aversion component $\left\{\varphi_3 * E[\tilde{I}_t|H_{t-1}] + \varphi_4 E\left[\sigma^2_{p,t}|H_{t-1}\right]\right\}\psi_t/2(\varphi_\mathrm{t} + \theta_t)$. In line with the MRR model, the variance of price change can be written as:[177]

$$Var(\Delta p_t) = \underbrace{\sigma_\varepsilon^2 + 2\sigma_\xi^2}_{\substack{Public\ Information \\ and\ Discretness}} + \underbrace{(1-\rho^2)\theta_t^2}_{\substack{Asymmetric \\ Information}} + \underbrace{2\varphi_t\theta_t(1-\rho^2)}_{\substack{Interaction\ between \\ Asymmetric\ Information \\ and\ Transaction\ Costs}} + \underbrace{\varphi_t^2 + \varphi_{t-1}^2 - 2\rho\varphi_\mathrm{t}\varphi_\mathrm{t-1}}_{\substack{Transaction \\ Costs}}. \tag{6.27}$$

Similar to the MRR model, this formulation decomposes price change variance into five components; variations due to Public Information, Price Discreteness, Asymmetric Information, Transaction Costs and Interaction between Asymmetric Information and Transaction Costs. Considering that information and liquidity components are functions of dealers' expectations, which in turn are functions of trading intensity and risk, the variance is, consequently, modelled as a function of expected trading intensity and risk. This allows past information to have an indirect impact on variance. This is a major extension to the MRR model, since return variance is recognized to vary due to the potential effects of the actual events in shaping future expectations, and thus learning. The impact of these expectations might be different in different environments, due to structural reasons, such as the nature of the underlying asset or the legal framework of the market. This is recognized by the dynamic character of the return model, Eqs. (6.22) and (6.23), which allows for a different formulation for expectations in each data set employed. Finally, similar to the MRR model, autocorrelation of the order flow, ρ, has a decreasing effect on variance. This accounts for the Bid-Ask bounce, defined as rapid price changes between Bid and Ask, when the price midpoint does not necessarily change, which is higher when there are no market trends, such as trading in the same direction, due to either information or liquidity needs.

Estimation and Further Applications of the Dynamic Model

According to the literature, the preferred method of estimating trade indicator models is the Generalized Method of Moments (GMM). In this method, the selection of appropriate orthogonality (i.e., population moment conditions that are assumed to be zero) conditions is essential. In this study, the estimation procedure follows the example

[177] A formulation similar to Madhavan et al.'s (1997) could be written as:
$Var\ (\Delta p_t) = \sigma_\varepsilon^2 + 2\sigma_\xi^2 + \{(\theta_t + \varphi_t)^2 + (\rho\theta_t + \varphi_{t-1})^2 - 2\rho(\theta_t + \varphi_t)(\rho\theta_t + \varphi_{t-1})\}$.
A full derivation is found in the chapter's appendix.

of Grammig (2002), Grammig et al (2007), Angelidis and Benos (2009) and Ben Sita (2010).

First, let $\beta = \left(\theta_n, \varphi_n, c_m, \rho, \sigma_\varepsilon^2, \sigma_\xi^2\right)'$, where $n = 1, \dots, 4$ and $m = 1, \dots, 6$, be a vector of the parameters to be estimated, υ_t a vector of all available variables at time t, $z_t^\upsilon = (s_t, r_t, q_t, I_t, D_t)'$ a vector of the explanatory variables of trading intensity and $z_t = (q_t, q_{t-1})'$ a vector that contains the sign of the current and preceding. In addition, let $D_t = \left(E[s_t|H_{t-1}], I_{\pi,t}, E[\tilde{I}_t|H_{t-1}], E\left[\sigma_{p,t}^2|H_{t-1}\right]\right)'$ be a vector of the formulated expectations of trading intensity, s_t, number of informed traders, $\tilde{I}_t$, and variance, $\sigma_{p,t}^2$, and of an indication of the regime, according to expected trading intensity, the next transaction is expected to be in, $I_{\pi,t}$, while $u_t = r_t - E[r_t|H_t]$.

For the modelling of the expected trading intensity, the following moment conditions can be implied.[178] First, the forecasting error, $\mathrm{e_t}$, $E[s_t|H_{t-1}]$ in (6.22), is assumed to have a zero mean ($E\left[f_t^i(\beta, \upsilon_t)\right] = E[\mathrm{e_t}] = 0$) and be uncorrelated ($E[f_t^c(\beta, \upsilon_t)] = E[e_t e_{t-1}] = 0$). Second, all independent lagged variables are assumed to be uncorrelated with $\mathrm{e_t}$ ($E[f_t^\upsilon(\beta, \upsilon_t)] = E[\mathrm{e_t} * z_{t-1}^\upsilon] = 0$).[179] Furthermore, for the return Eq. (6.23), a very important moment condition is related to the autocorrelation of order flow. This is, $E\left[f_t^{of}(\beta, \upsilon_t)\right] = E[\mathrm{Cov}(q_t, q_{t-1}) - \rho Var(q_t)] = 0$. In addition, Madhavan et al. (1997) define a drift term, a, using the following moment condition, ($E[f_t^d(\beta, \upsilon_t)] = E[\mathrm{u_t} - \mathrm{a}] = 0$). Then in order to estimated the variance of public information, σ_ε^2, and Price Discreteness, σ_ξ^2, disturbances, they use the following functions; ($E[f_t^{var}(\beta, \upsilon_t)] = E\left[(\mathrm{u_t} - \mathrm{a})^2 - \left(\sigma_\varepsilon^2 + 2\sigma_\xi^2\right)\right] = 0$) and ($E[f_t^{ac}(\beta, \upsilon_t)] = E\left[(\mathrm{u_t u_{t-1}}) + \sigma_\xi^2\right] = 0$). Moreover, the remaining coefficients in Eq. (6.23) are decomposed employing the following restrictions $E[f_t^D(\beta, \upsilon_t)] = E[\mathrm{u_t} z_t] =$

[178] Moment conditions are functions, $E\left[f_t^\xi(\beta, \upsilon_t)\right]$where $\xi = i, c, \upsilon, of, d, var, ac, D$, derived from the estimated model, that include the parameters, β, to be estimated and variables, υ_t, observable at time t, and they are assumed to be zero. GMM estimates the model parameters, β, minimizing the sample averages of these moment conditions to be as close as possible to the (zero) population mean.

[179] Wooldridge (2001) mentions that the most important departure from common OLS assumptions, in the application of linear time series, is that the errors are serially correlated. Hansen (1982), White and Domowitz (1984) and Newey and West (1987) argue that implying over-identifying restrictions allows the GMM weighting matrix to account for serial correlation of unknown form. This can be done by including lagged variables. In case extra moment conditions are needed, then according to Grammig and Wellner (2002) the moment condition that the forecasting error is uncorrelated with each explanatory variable can take the form; $E\left[f_t^\upsilon(\beta, \upsilon_t)\right] = E\left[e_t * (z^\upsilon)_{t-j}^l\right] = 0, for\ t > 1, l = 0,1\ and\ j = 1, \dots, +\infty$. It follows that, depending on the appropriate lag length employed by the researcher, the autocorrelation restriction can be further extended, as in Grammig et al. (2007). $\mathrm{E}[\mathrm{f_t^c}(\beta, \upsilon_t)] = \mathrm{E}\left[\mathrm{e_t} * \mathrm{e_{t-j}}\right] = 0\ for\ j = 1, \dots, +\infty$. Furthermore, the inclusion of r, as instrumental variable in $E[s_t|H_{t-1}]$, accounts for the interdependence between trading intensity and return disturbances.

$E[\mathrm{u_t}z_tD_t] = E[\mathrm{u_t}z_{t-1}D_{t-1}] = 0$. Summarizing, the joint model is estimated using the following moment conditions:

$$E\left\{\begin{array}{c} \mathrm{e_t} \\ e_te_{t-1} \\ \mathrm{e_t}z_{t-1}^{\upsilon} \\ q_t, q_{t-1} - \rho q_t^2 \\ (\mathrm{u_t} - \mathrm{a}) \\ (\mathrm{u_t} - \mathrm{a})\mathrm{z_t} \\ (\mathrm{u_t} - \mathrm{a})\mathrm{q_t}\mathrm{D_t} \\ (\mathrm{u_t} - \mathrm{a})\mathrm{q_{t-1}}\mathrm{D_{t-1}} \\ (\mathrm{u_t} - \mathrm{a})^2 - \left(\sigma_\varepsilon^2 + 2\sigma_\xi^2\right) \\ (\mathrm{u_t}\mathrm{u_{t-1}}) + \sigma_\xi^2 \end{array}\right\} = 0. \tag{6.28}$$

The GMM disturbances are then gathered in a vector $(\beta, \upsilon_t) = [f_t^i(\beta, \upsilon_t)', f_t^{\upsilon}(\beta, \upsilon_t)', f_t^{of}(\beta, \upsilon_t)', f_t^d(\beta, \upsilon_t)', f_t^{var}(\beta, \upsilon_t)', f_t^{ac}(\beta, \upsilon_t)', f_t^D(\beta, \upsilon_t)']'$, and the sample means are defined as; $g(\beta; S_T) = \frac{1}{T}\sum_{t=1}^{T} f(\beta, \upsilon_t)$, where S_T contains the observations of $\upsilon_{t-j}, j = 0, \dots, T$ of a sample *T*.

The idea behind GMM is to choose parameter values for β, such that the sample moments, $g(\beta; S_T)$, closely approximate the population moments, $f(\beta, \upsilon_t)$, or else to make $g(\beta; S_T)$ as close to zero as possible. By the "*Law of Large Numbers"* $g(\beta; S_T) \approx f(\beta, \upsilon_t)$ for large values of *T*, so an appropriate estimate, $\hat{\beta}$, of the population parameter β_o makes $g(\beta; S_T) \approx 0$. When the number of moment conditions, *K*, is larger than the number of parameters, *L*, then the GMM estimator can be written as:

$$\hat{\beta} = arg\min_{\beta}\left(g(\beta; S_T)' * \widehat{W}_t * g(\beta; S_T)\right), \tag{6.29}$$

where $\widehat{W}_t$ is a KxK semi-definite "weighting" matrix, such as that $\lim_{T\to\infty} \widehat{W}_t \to W$ (population).[180] The approach employed for the estimation of $\hat{\beta}$ is the "*iterative*" GMM, with a heteroskedasticity consistent covariance matrix (Newey and West, 1987), when computing an estimate for $\widehat{W}_t = \Omega^{-1}$ (where $\Omega = \lim_{T\to\infty} T * E[g(\beta; S_T) * g(\beta; S_T)']$).

In the above specification, $K > L$ and, therefore, the model is over-identified. Hansen (1982) proposes *J-statistics* to test the validity of the model, i.e., whether the implied

[180] Hansen (1982) shows that $\hat{\beta}$ is consistent, asymptotically normal and, with the right choice of $\widehat{\mathrm{W}}_{\mathrm{t}}$, asymptotically efficient.

moment conditions fit the data well. H_o is that they do. *J-statistic* is asymptotically *Chi-squared* with $K - L$ degrees of freedom.

$$J \equiv \left(g(\beta; S_T)' * \widehat{W}_t * g(\beta; S_T) \right) \rightarrow \chi^2_{K-L}. \tag{6.30}$$

Besides testing model validity, the in-sample and out-of-sample performance of the MRR model and the extensions will be compared, by the use of the Mean Squared Error (MSE).

Another important issue that is examined is whether the enhanced MRR model can help traders in deciding which type of order to submit; market or limit order. The basic assumption here is that the information aggregated by previous transactions can formulate an expectation about the informational content of the next trade. The model describes two major forces to be the drivers of price formation: information, which is summarized by θ, and liquidity, which is summarized by φ. Information moves prices permanently to a new equilibrium, while liquidity considerations have only a transitory impact. Suppose that trading intensity has an impact on both of them, the magnitude and the direction of this impact could be of critical importance. If there are trades that are expected to increase θ more than φ, or even decrease φ, then these trades are more likely to move prices permanently. Therefore, investors' wealth would benefit from the submission of an appropriate market order, according to the expected direction of the new equilibrium. However, if the permanent price impact of a transaction is expected to be low, i.e., lower than the transitory component, investors can earn the spread, instead of treating it as an implicit trading cost, by submitting a limit order. Consequently, although quite rigid, the joint model can offer a tool in selecting the most beneficial order type.[181] If the next transaction is expected to have a high information impact on price, then investors should aim at capturing capital gains after paying the spread, but if a low permanent price impact is expected, then investors should aim at providing liquidity and earn the spread, obviously while bearing the risk of return volatility.

One way to examine the expected information and liquidity content of trades is to work at basic statistics of θ_t and φ_t. Their average values are calculated across various levels of expected trading intensity, price volatility, level of OTC and informed transactions,

[181] Trading intensity is assumed to be the only determinant of the informational content of the next trade. This provides a link with the material presented in the previous chapter. However, other variables such as price volatility and OTC or informed transactions are recognized to have a price impact, and even though they are not modelled directly here, their effect is empirically examined, through basic statistics.

as well as during the trading day. Larger, and probably increasing, θ_t would be an indication that market orders could be more profitable, especially when it is associated with lower values of φ_t. In the opposite case, a limit order strategy should be preferable.

Another way to measure the appropriateness of the type of order submission, while accounting for the sensitivity of the permanent and transitory price impact of trades, is to compare the rate of change of return variance with the rate of change of the implied spread for a given change in expected trading intensity. This measures the sensitivity of variance and spread towards expectations, and can be calculated by the partial derivatives with respect to expected trading intensity. This is:

$$R_t = \frac{\partial[Var(\Delta p_t)]}{\partial E[s_t|H_{t-1}]} - \frac{\partial\left[S_t^{Implied}\right]}{\partial E[s_t|H_{t-1}]} \lesseqgtr 0 \tag{6.31}$$

where R_t is the difference between the rate of change of the implied spread, which is computed as $2\left\{\sum_{\pi}^{3}\left[\left(\theta_{2,\pi} + \varphi_{2,\pi}\right)I_{\pi,t}\right]\right\}$ and the rate of change of the return variance (the full derivation is given in the appendix). When $R_t > 0$, the rate of change of variance is expected to be higher than the rate of change of the implied spread. In this case, the risk of exposure is expected to be higher than the compensation paid by the implied spread and, therefore, no market making can be considered profitable. Conversely, when $R_t < 0$, the change in the implied spread due to variations in expectations is higher than that of the variance, and, therefore, limit orders should be considered more appropriate, since they are less risky. If the above analysis provides evidence of a consistent pattern, then this would lead to the conclusion that managing both transaction size and time should be profitable, or at least less expensive.

6.3 Empirical Results

6.3.1 Estimation

Table 6.1, which consists of the Tables A. ECX I, B. ECX II, C. NP I and D. NP II, reports the estimation results for the original, MRR, and the Dynamic Expectations Joint models, for all markets and phases. In more detail, the first column presents the estimates of the MRR model. In the next three columns, the coefficient theta, θ_t, is decomposed, as in Eq. (6.16), following a stepwise regression analysis. First, $\theta_{2,\pi}$ coefficients are added, which are related to expected trading intensity, $E[s_t|H_{t-1}]$, T.I., while in the middle column, Risk, the risk factors, captured by the coefficients θ_3 and θ_4, are examined separately. Then their combined impact on theta is examined in the last column, named T.I + Risk. Along the same lines, the following three columns decompose phi, φ_t, into T.I, Risk and their combined effect, according to Eq. (6.18). The last two columns present the full model, where both θ_t and φ_t are decomposed, with the last one including only the statistically significant parameters. Furthermore, the first six rows decompose θ_t across models, while the next six decompose φ_t. Then the autocorrelation of the order flow, ρ, along with the coefficients of Eq. (6.12) and the variances of the innovations, σ_ε^2 and σ_ξ^2, are presented. *t-statistics* are in parentheses. The last two sections present the *J-statistics* and the associated *p-values* in parentheses, as well as the Mean Squared Error (MSE) for in- and out-of-sample forecasts.

The MRR model

The first point to note is that the parameter estimates of the MRR model do not significantly differ in sign and magnitude from the estimates of the same model applied to other trading environments such as the New York Stock Exchange (Madhavan et al.,1997) and the Athens Stock Exchange (Angelidis and Benos, 2009). They are also consistent with those reported by Benz and Hengelbrock (2008) and Frino (2010) for European Climate Exchange (ECX) and Nord Pool (NP). More specifically, as can be seen in the first column of Tables 1.A, 1.B, 1.C and 1.D, the information component, θ, is considerably larger (0.0808 in ECX I, 0.0496 in ECX II, 0.1201 in NP I and 0.0657 in NP II) than the transaction cost component, φ, (0.0410 in ECX I, 0.0182 in ECX II, 0.0369 in NP I and 0.0597 in NP II). Both of them are statistically significant in both markets and phases, but θ's *t-statistics* are consistently much higher (i.e., in ECX it is almost three times larger in Phase I (9.31 and 3.26) and almost five times larger in Phase II (15.38 and 3.10)).[182] This is an initial sign indicating that information is the driving force in intraday formation in the Carbon market.

Furthermore, the sum of θ and φ, which according to the MRR model is half spread, is consistently larger in Phase I ($\theta + \varphi = 0.1218$ in ECX and 0.1570 in NP) compared to

[182] This opposes Benz and Hengelbrock (2008) who report very small and, sometimes, insignificant estimates for the transitory component. However, they estimate the MRR model dividing their data sets per Quarter of the Calendar year. This seems to have a decreasing impact on the significance of the cost component, which seems to reduce in Phase II and especially in Nord Pool. In this study, since φ and θ are decomposed into continuously updated components, the model is estimated once for the whole period. Specific values for given periods of time can then be computed using basic statistics. This seems to boost the significance of the parameters (at least on average). Finally, another, critical difference is that they exclude OTC transactions, which are expected to have a statistically significant price impact.

Phase II (0.0678 in ECX II and 0.1254 in NP II), especially in the less liquid market, Nord Pool. In addition, the information component decreases significantly in Phase II in both markets (from 0.0808 to 0.0496 in ECX and from 0.1201 to 0.0657 in NP). Another similarity with results on NYSE is the highly significant presence of autocorrelation in order flow, which seems to be substantially higher in ECX (0.5022 in Phase I and 0.4756 in Phase II) than in NP (0.2708 in Phase I and 0.2946 in Phase II). This is consistent with the analysis in Chapter 5, Figure 5.2.C, where the autocorrelation of order flow is consistently lower in NP, raising considerations about its connection with liquidity. Furthermore, the variances of the public information and Price discreteness are significantly higher in Phase I (0.0066 and 0.00017 in ECX and 0.0046 and 0.00025 in NP) than in Phase II (0.0034 and 0.00005 in ECX and 0.0018 and 0.00016 in NP)), indicating a less volatile environment as the market grows and matures.

This preliminary analysis provides some initial evidence concerning the evolution of intraday formation of prices and trading spreads in the Carbon market. The narrower spreads that can be observed in Phase II, indicate a progress to maturity, mainly due to a decrease in the information component and they raise liquidity considerations. Given the increased liquidity in both markets, reported in Chapter 3, Figure 3.5 and in Chapter 5, Table 5.6, this can be interpreted as a decrease in the proportion of informed traders or an indication that a lower compensation dealers need for transacting with them, or both. Figures 5.1.A and 5.1.B show clearly a decrease in the proportion of information based trading in Phase II. This could be the result of either an increase in uninformed transactions or the provisions of less information incentives, as a result of improved information dissemination. Therefore, dealers seem to decrease the information component of the spread either because the risk of transacting with an informed trader is lower or because they can earn more by transacting with uninformed traders. In addition, spreads appear to be narrower in the largest and more liquid market (ECX), underlining the importance of market size and depth. Finally, what is also noticeable is that although the public innovations' variance is consistently larger in ECX, the variance of price discreteness is consistently higher in NP. This could be an initial indication that ECX is the price leader and more mature. Information shocks are higher, while, pricing is more accurate (i.e., lower price discreteness variance).

Theta and Phi

The next two sections of Table 6.1 decompose theta, θ, and phi, φ, in order to take into account trading intensity, T.I., and risk, Risk, expectations, as discussed in Section 6.2.2. The first three columns under theta decompose the information component, θ_t, as in Eq. (6.16). A close inspection reveals a decrease in magnitude and significance of the initial parameter θ, of the MRR model (0.0808 to 0.07771 in ECX I, 0.0496 to 0.0326 in ECX II, 0.1201 to 0.0285 in NP I and 0.0657 to 0.0508 in NP II), when expected trading intensity is taken into account. In addition, the significance of that parameter decreases dramatically (9.31 to 2.06 in ECX I, 15.38 to 2.02 in ECX II) and in some cases it becomes insignificant (6.95 to 1.75 in NP and 2.38 to 1.56 in NP II), indicating that expected trading intensity could be an important determinant of intraday price variations. Risk, in contrast, fails to provide any statistically significant change.

A closer inspection of the impact of expected trading intensity on the information component, in the columns under T.I. and T.I+Risk, reveals that it increases theta across all different regimes, although with a different rate. This is consistent with Dufour and Engle (2000b), in the sense that more frequent trading carries higher informational content and, therefore, has a higher price impact. The same seems to happen for larger transaction size as well, which confirms Easley and O’Hara (1987), Hausman et al. (1992), Easley et al. (1997) and Angelidis and Benos (2009), who argue that higher trading volume indicates the presence of informed trading and therefore that these trades should have a higher price impact. The regimes of fundamental and informed trades in particular have always a statistically significant effect, with the informed regime having consistently higher *t-statistics*. In addition, in ECX when trading intensity is included along with theta, the magnitude and the significance of phi decrease as well. This emphasizes the role of size and time on determining the price impact of the adverse-selection component. In contrast, in Nord Pool, especially in Phase I, theta becomes insignificant (*t-stat* = 1.75) and phi plays a more significant role in pricing, since it increases in both magnitude (0.0369 to 0.0592) and significance (2.01 to 3.64).

The next section of Table 6.1, under phi, decomposes the liquidity component φ, according to Eq. (6.18) into the impact of expected trading intensity, T.I., and risk, Risk. When expected trading intensity is taken into account, a major change in magnitude (0.0410 to 0.0760 in ECX I, 0.0182 to 0.0308 in ECX II, 0.0369 to 0.0491 in NP I and 0.0597 to 0.0700 in NP II) is observed in the parameter φ of the MRR model. In addition, the impact of expected trading intensity appears to be statistically significant, especially in the fundamental and the informed traders' regimes. Coefficients $\varphi_{2,\pi}$ are all negative and mostly statistically significant (e.g., in ECX II $\varphi_{2,uninformed}$ is -0.0117 (-2.54), $\varphi_{2,fundamental}$ is -0.0109 (-2.93) and $\varphi_{2,informed}$ is -0.0035 (-2.97)) especially in the fundamental and informed regimes.

This suggests a negative relation between prices and expectations of trading intensity. This negative relation causes φ_1, which is a long-term average, to be much larger, and this is consistent with Angelidis and Benos (2009). They mention that due to economies of scale, transaction size is expected to have a decreasing price effect. Moreover, this relation seems to be enhanced by time as well, since expected trading intensity consists of size and duration. Higher trading frequency reduces trading costs in an analogous fashion to transaction size. Consequently, in accordance with previous literature, it seems that a fixed cost (captured by φ_1) exists. This cost decreases as trading intensity increases. This is probably due to economies of scale, leading to more opportunities to trade or more efficient cost allocation.

In addition, unlike theta, risk expectations appear to be significant determinants of the liquidity component. However, they are summarized in price volatility, as informed trading fails to provide statistically significant results. Coefficient φ_4 is always positive, and statistically significant, indicating that spread increases when dealers, or limit order traders, expect to be exposed to higher price volatility for longer.

Several comments are relevant to these findings. First, the transitory price component is found to be inversely related to trading intensity expectations, compared to the information component. Ben Sita (2010) report a transitory and a permanent pricing component related to the time dimension of trades. Bowe et al. (2007) extend this idea to include transaction size, which in Angelidis and Benos (2009) is related to permanent and transitory effects. The present study confirms the empirical findings of these studies and summarizes the effects of both on trading intensity. Consequently, trading intensity seems to play a dual role in price formation, affecting in a different way information and liquidity.

This becomes more obvious by looking at Table 6.2, which presents the average expected trading intensity and the estimates of theta and phi, along with half spreads, according to the estimates of the last columns of Table 6.1. More specifically, the first column, Average Trading Intensity, presents the average, arithmetic mean, of the expected trading intensity in each regime, in each phase and in each market. The next column, Theta, presents the estimates of theta across regimes, markets and market phases, while the third column, Average Theta, is the product of the previous two and indicates the average adverse-selection component for each regime in the different datasets. Along the same lines, the next two columns present the estimates of the liquidity component, phi, and its average values.[183] The last column is the sum of Average Theta and Average Phi and is half the implied spread.

The Average Theta increases across different regimes of expected trading intensity (e.g., 0.0533 in the uninformed regime, 0.0791 in the fundamental and 0.0977 in the informed in ECX I). This can be seen as the need for higher compensation due to a higher presence of informed traders. When dealers or limit order traders expect higher transaction size and/or higher trading frequency, they see that as a signal of more prominent presence of informed traders. Therefore they widen the spread to compensate for potential losses. In contrast, Average phi follows a U-shape pattern (0.0489 in the uninformed regime, 0.0060 in the fundamental and 0.0657 in the informed in ECX I). This could indicate that, in practice, dealers manage their Inventories, in a similar fashion to “inventory” models (inter alia, Stoll, 1978), and that they have an optimal position that they want to return to, either because it is less risky, or less costly. Consequently, when trading intensity is expected to be low they charge more for providing liquidity. When it is high they widen the spread, so they will not deviate significantly from their target. In between, there should be a range where inventory deviations, as well as “immediacy” costs, are not that high, and thus there is a lower liquidity cost component. This seems to explain the differences between Stoll (1978), who describes a U-shape relation between spread and transaction size, De Jong et al. (1995, 1996) who maintain that spread is a decreasing function of volume, and Huang and Stoll (1997) and Ahn et al. (2002), who maintain that order-processing cost is a decreasing function of volume. Trading intensity is negatively related to the liquidity component, but that decreasing relation varies in magnitude in different regimes and thus the U-shape relation.

The MRR model seems to summarize a large order-processing cost component, and its adjustments to liquidity and risk variations, in one parameter. According to the results

[183] For clarity, Table 6.2 emphasizes the impact of Expected Trading Intensity on half spreads and, therefore, Risk is not present in the calculations.

that seems to be quite a rigid approximation and it probably underestimates this component, unlike Huang and Stoll (1997) who report a much higher proportion of order-processing cost. This might also be the reason why Benz and Hengelbrock (2008) report a small and insignificant transaction cost component, φ, after estimating the MRR model in the Carbon market.

Full Model

The last two columns of Table 6.1 present the estimation results of the full model, as in Eqs. (6.22 and 6.23). The last column especially, named Full-Significant, presents the final version of the joint model that includes only the statistically significant parameters. Focusing on Eq. (6.23) and on the components of the spread, the only significant parameters of θ_t are the ones associated with expected trading intensity, while for φ_t only the expected Informed Trading is not relevant. Therefore, unlike phi, theta seems to be directly proportional to trading intensity, without any long-term average, since θ_1 is statistically insignificant (*t-stat* is 0.35 in ECX I, -0.40 in ECX II, -1.47 in NP I and -0.95 in NP II). This is consistent with Angelidis and Benos (2009), who report that the adverse-selection component is an increasing function of trade volume without a long-term average. In contrast, phi can be further decomposed into order-processing, φ_1, liquidity, $\varphi_{2,\pi}$, and risk aversion, φ_4, components. Consequently, according to this model, prices and spreads are the outcome of balancing expectations of the dual effect of trading intensity and risk.

Another important comment concerns the trading intensity equation, presented in Eq. (6.22). Trading intensity appears to be autoregressive, since c_2 is statistically significant (*t-stat* is 15.24 in ECX I, 23.39 in ECX II, 7.51 in NP I and 3.10 in NP II), but, surprisingly, it seems that it is not endogenously related to returns (e.g., *t-stat* is -0.14 in ECX I, -0.46 in ECX II) or order flow (e.g., *t-stat* is 0.92 in ECX I, 1.84 in ECX II). In contrast, a negative (-1.2069 in ECX I, -0.4697 in ECX II, -1.2723 in NP I and -1.1223 in NP II) and statistically significant (*t-stat* is -28.78 in ECX I, -14.60 in ECX II, -6.72 in NP I and -6.41 in NP II) estimate of parameter c_6, indicates that OTC transactions have a statistically significant decreasing effect on expectations. This is consistent with the empirical finding in Chapter 4 (Eq. (4.41) and Tables 4.3.B, 4.4B, 4.5.B and 4.6.B), where $\zeta > 0$, and OTC transactions increase the expected duration and therefore decrease the expected trading intensity.

A positive (0.2271 in ECX II, 0.5765 in NP II) and statistically significant (*t-stat* 6.53 in ECX II and 2.76 in NP II) estimate of parameter c_5, as in Eq. (6.12), indicates that as the market gains complexity over time, from Phase I to Phase II, informed transactions play a more crucial role in formulating expectations of trading intensity. This might happen either because informed traders act strategically in Phase II, or because of more efficient information dissemination. In the former case, one informed transaction is likely to be followed by similar transactions (strategic), which should increase trade frequency, and, therefore, the traded contracts per unit of time. In the latter case, price-relevant, unresolved information is more efficiently disseminated and market participants, such as dealers, who observe order flow to extract price-relevant information, can track it and benefit from the time space till it goes public. Trading intensity is expected to increase in that period of time.

Empirical evidence provided in Chapter 5 tends to support that the most relevant scenario is mostly determined by the market. More specifically, Figure 5.2.B shows that the average transaction size increases in Phase II, while Figure 5.3 shows that the average duration consistently decreases. This indicates a general increase in trading intensity, which, as Figures 5.1.A and 5.1.B suggest, is not necessarily being followed by an analogous increase in informed trading, since the proportion of informed trades is decreasing in Phase II. Furthermore, Figure 5.2.C shows that the autocorrelation of the order flow in the informed regime is lower in ECX II, compared to ECX I, but higher in NP II, compared to NP I. Likewise, Table 5.7 shows that the proportion of identical trades in informed regime, which would indicate strategic trading, is lower in ECX II (19.51 percent), but higher in NP II (14.29 percent), compared to Phase I (31.77 percent and 0 percent respectively). This suggests that in Phase II, the actions of informed trades are more likely to be revealed in ECX II, while informed traders appear to act strategically in NP II. In both cases their actions increase the expected trading intensity. Further, the sign of better information dissemination is also another indication that ECX, probably due to its higher liquidity, is a more efficient market compared to Nord Pool.

Finally, the last two rows of Table 6.1 present the performance of all estimated models according to the Mean Squared Error (MSE). MSE seems to reward the enhanced flexibility of the joint model, especially, when only the statistical significant parameters are taken into account. It is consistently smaller in magnitude across markets and phases in both in- and out-of-sample forecasts. This indicates that the original, MRR, model is restrictive, because it ignores trading intensity variations.

6.3.2 Spread as continuous variable

Spread and Variance Components

One of the characteristics of the model, described in Section 6.2.2 in Eqs. (6.22) and (6.23), refers to describing the Bid-Ask spread as a continuous variable. Expectations of theta and phi are continuously revised after every transaction, according to Eqs. (6.16) and (6.18). This means that on the occurrence of a transaction, market participants are assumed to formulate expectations concerning the incoming trading intensity and the price volatility. Using these expectations they determine the compensation they require for the probability of trading with better informed traders and for supplying liquidity, and they revise the Bid and Ask prices they are willing to trade, and thus the spread. Consequently, any intraday pattern of the spread or the price change variance, as well as their components, can be examined through basic statistics of θ_t and φ_t, according to Eqs. (6.24), (6.25) and (6.27), without re-estimating the model. Table 6.3, from which Figures 6.1, 6.2, 6.3, 6.9, 6.10, 6.11, 6.17, 6.18, 6.19, 6.25, 6.26 and 6.27 are derived, presents the intraday formation of spread, spread components, price change variance and its components.

Specifically, the first three columns of Table 6.3 present the Time of the Day, the Trading Intensity and the Expected Trading Intensity, as it is formulated by Eq. (6.22) and the estimates in the last column of Table 6.1. The next section presents the Adverse-selection component, which is θ_t as in Eq. (6.16), and Phi, which is the sum of Liquidity, $\varphi_1 + \sum_{\pi}^{3}\left(\varphi_{2,\pi} I_{\pi,t}\right) E[s_t|H_{t-1}]$, and Risk Aversion, $\left\{\varphi_4 E\left[\sigma_{p,t}^2|H_{t-1}\right]\right\} * \psi_t$, as in Eq. (6.18). The next section presents the Implied and Effective Spread, which is double the sum of Adverse-selection and Phi, as in Eqs. (6.24) and (6.25). The next section presents the Asymmetric Information, Trading Costs and Interaction, between them, components of the price change variance, as in Eq. (6.27), which is presented in the last column. Furthermore, panel A in figures 6.1, 6.9, 6.17 and 6.25 presents graphically the two sections of Table 6.3, that refer to the Trading Intensity, the Expected Trading Intensity and the Bid-Ask Spread, Implied and Effective, while panel B presents the section that refers to the spread components. Panel C summarizes intraday variations of variance and its components. Next, panels A and B of Figures 6.2, 6.10, 6.18, 6.26 decompose the spread in its components, in absolute numbers and as percentages, while panels C and D do the same for variance. Finally, Figures 6.3, 6.11, 6.19 and 6.27 show the intraday variation of the relation between spread, in panel A, or variance, in panel B, and expected trading intensity.

Looking at the columns that refer to the Implied and Effective Spread in Table 6.1, and in panel A of Figures 6.1, 6.9, 6.17 and 6.25, Spread consistently follows a U-shape pattern. In some cases, as in ECX, a tube-shaped pattern better describes the spread variations, because of a relatively long period with small fluctuations. In the beginning of the trading day, for the first hour the spread is wide and it starts declining. It stays quite stable, with an exception around lunch break where it increases marginally, up until an hour before the official closing of the market, where it increases again. During Phase I the spread at the beginning of the day (the average spread is 25.73 cents in ECX and 38.8 cents in NP) is larger than at the closing, but during Phase II the closing session appears to be more volatile and, therefore, wider spreads are observed (from 14.87 to 16.53 cents in ECX and from 53.92 to 54.59 cents in NP, in the opening and closing sections, respectively). Narrower spreads are observed just after the opening or before the closing sessions. A close inspection of the graphs reveals that fluctuations in spreads follow the fluctuations in trading intensity, which appears to be a major determinant of these changes. Trading Intensity, and expected trading intensity, are either decreasing, in Phase I, or U-shaped, in Phase II, but always declines sharply over the first trading hour and increases over the last, while the lowest point is always around midday, the well-known lunch break effect. This is when a slight increase in the spreads is observed. Consequently, the level of expected trading intensity appears to influence the Bid-Ask spread, either because the information component or the cost of providing liquidity, or both, are high.

This becomes more obvious by examining panel B, 6.1, 6.9, 6.17 and 6.25, along with the section of Table 6.3 that refers to spread components. Spread components follow patterns similar to the ones reported in Madhavan et al.'s (1997) original paper. The asymmetric information component, θ_t, described in Eq. (6.16), generally declines over the day (e.g., in ECX I θ_t is 7.52 cents during the opening, while the lowest point is at closing at 5.20 cents), with a slight increase in the closing session (e.g., in NP II θ_t is 11.40 cents during the opening session, it decreases to 6.03 cents an hour before closing and increases slightly, to 8.14 cents, during the closing session) or it exhibits a U-shape pattern (e.g., in NP I θ_t is 15.6 cents during opening, the lowest point is around lunch time at 11.50 cents, while at closing increases to an average of 13.75 cents). A comparison between panels A and B in the above figures indicates that θ_t variations seem to follow, in a rigid way, the fluctuations of trading intensity. This is to be expected since θ_t is directly proportional to expected trading intensity, according to the last column in Table 6.1. Moreover, according to panel B of Figures 6.2, 6.10, 6.18 and 6.26, adverse-selection accounts for 40 to 80 percent of the total spread, depending on the market, the phase and the time of the day.

In contrast, the liquidity component, as it is summarized in coefficients, φ_1 and $\varphi_{2,\pi}$ in Eq. (6.18), follows the opposite pattern and accounts for 30 to 50 percent of the total spread. It is generally increasing during the trading day and the peak is always during the closing session (6.62 cents in ECX I, 2.57 cents in ECX II, 4.01 cents in NP I and 6.37 cents in NP II). The lowest level in Phase II is observed during the opening session (1.44 cents in ECX and 5.24 cents in NP), while in Phase I it is one hour later (4.33 cents in ECX 2.57 cents NP). The risk aversion component, which according to the last column in Table 6.1 is summarized in coefficient φ_4, Eq. (6.18), in ECX and NP I, and in coefficient φ_3 in NP II, generally increases over the trading day, although it is small in magnitude. The maximum values are always observed during the closing session (0.93 cents in ECX I, 0.68 cents in ECX II, 1.38 cents in NP I and 3.24 cents in NP II), while this component accounts, as can be seen in panel B of Figures 6.2, 6.10, 6.18 and 6.26, for less than five percent of the total spread. Looking at panels A and B of Figures 6.1, 6.9, 6.17 and 6.25, that is sufficient to explain part of spread variations, especially when expected trading intensity is in a middle range.

Moreover, considering that theta and phi, as they are defined in Eqs. (6.16) and (6.18) respectively, measure price sensitivities, these variations are to be expected. At the beginning of the trading day there is overnight information that is yet to be observed. Therefore, dealers' reactions, and consequently their quoted prices, are expected to be very sensitive to trades, in case they carry information. However, as time lapses this information is incorporated into prices and dealers aggregate data as well, and the information component, θ_t, declines. The revived trading activity at the end of the day increases, again, the probability of informed trading and therefore the information component. In contrast, following the U-shaped formation of φ_t (excluding Risk Aversion) across expected trading intensity, presented in Table 6.2, the liquidity component, $\varphi_1 + \sum_{\pi}^{3}\left(\varphi_{2,\pi} I_{\pi,t}\right) E[s_t|H_{t-1}]$, seems to depend on the configuration of dealers' portfolios. They appear to have an optimal inventory position, which is affected by trading intensity. Consequently, different levels of trading intensity should affect their actions in a different way, and this should result in a preferred range of market activity that minimizes risks or costs. Therefore, assuming that their preferences are not time variant, the liquidity component should be expected to increase over the trading day, since there is less time to adjust to new information or liquidity needs. Along the same lines, the risk aversion component, $\varphi_3$or φ_4, should be expected to increase over the trading day for similar reasons.

Panel C of Figures 6.1, 6.9, 6.17 and 6.25 presents the intraday pattern variance of price change, $Var(\Delta p_t)$, and its components in ECX I, ECX II, NP I and NP II, respectively. As has been explained in Section 6.2.2 in Eq. (6.27), the variance consists of public information shocks, σ_{ε}^2, rounding errors, σ_{ξ}^2, and a trading component. Since the variance of the error terms has been implicitly assumed to be constant, any intraday variations are attributed to asymmetric information, $(1-\rho^2)\theta_t^2$, trading costs fluctuations, $2\varphi_t\theta_t(1-\rho^2)$, and any interaction between them, $\varphi_t^2+\varphi_{t-1}^2-2\rho\varphi_{\mathrm{t}}\varphi_{\mathrm{t-1}}$.[184] Consequently the variance is expected to exhibit similar patterns to spreads, which are also determined by the same components. Indeed, it exhibits the familiar U-shaped intraday variation, where it appears to be higher in the opening and the closing sessions and around the lunch break. In more detail, in Phase I both markets appear to be volatile during the opening session (0.0206 in ECX I and 0.0408 in NP I) whereas in Phase II, maximum volatility is observed during the closing session (0.077 in ECX II and 0.0420 in NP II). The minimum variance is usually observed just before the closing session (0.0193 in ECX I, 0.0065 in ECX II and 0.0310 in NP II) and it usually follows a local maximum (0.0199 in ECX I, 0.0068 in ECX II and 0.0363 in NP II), which is observed around the lunch break.

A large information component at the beginning of the trading day (0.0061 in ECX I, 0.0023 in ECX II, 0.0226 in NP I and 0.0148 in NP II) progressively decreases in Phase I, while in Phase II it exhibits an increase during the closing section. In contrast, the trading costs component is low during the opening (0.0028 in ECX I, 0.0004 in ECX II, 0.0021-0.0016 in NP I and 0.0066 in NP II) and progressively increases during the

[184] In Chapter 5, information shocks have been assumed to arrive at the market randomly, with frequency of arrival that follows a Poisson distribution. This does not allow for any systematic intraday variation in its variance. Therefore, for consistency, a constant variance is assumed throughout the period of each data set. In addition, even in the MRR model, rounding error variance is found to be very small and mostly insignificant. In the model presented in Eqs. (6.22) and (6.23), the variance is restricted to be constant, but any potential loss of information, because of this generalization, is expected to be limited. On the contrary, the difference between Phase I and Phase II is expected to be substantial because of market improvement and thus a new pair of coefficients σ_{ε}^2 and σ_{ξ}^2 will be estimated in each market.

trading day. These opposing intraday variations could happen because information is resolved over the trading day and, therefore, dealers do not need to charge more, which would result in higher θ_t, to compensate for transacting with better informed traders, while liquidity considerations rise because of the limited time they have to react closer to the closing of the market. Their movements are in the opposite direction and their effect seems to cancel out. Moreover, the interaction component, which measures their combined outcome, seems to have a major impact on variance, since they seem, with the exemption of ECX I, to follow similar patterns.

Along the same lines, Figures 6.2, 6.10, 6.18 and 6.26 as well as Figures 6.3, 6.11, 6.19 and 6.27 confirm the same findings. The first group describes intraday variations of half spreads and price change variance and their components in both absolute and relative terms for ECX I, ECX II, NP I and NP II. The second consists of “area” graphs that present visually the spread and return variance against expected trading intensity and time of the day, for the two markets and phases. Both spread and return variance exhibit a U-shape pattern both over time and across expected trading intensity. The lowest values are observed in the fundamental trading regime and especially after the opening and before the closing sessions. In addition, the public information contribution to the variance accounts for around 40 percent to 50 percent in ECX and 5 percent to 15 percent in Nord Pool. Given that ECX is the dominant market, in terms of size, this could be an indication that public information is first resolved in the largest market, which is expected to be the price leader.

Contrary to the MRR model, the trading frictions component in the Carbon market does not consistently increase over the day, and, consequently, the significance of public information does not decrease. In contrast, according to panel D of the first group of graphs, the proportion of return variance attributed to public information shocks increases during the day and it is lower during the opening and closing sessions, where the activity of the market is higher. In the opening and the closing sessions, when the most intensive trading is observed, the main determinant of return variance seems to be the formulated expectations about trading intensity and risk, and, therefore, the trading process itself. However, during the day, trading intensity calms down. Consequently, the contribution of the trading process to the variance is lower and price fluctuations are mainly caused by exogenous public information. This, together with the U-shape pattern of return variance, seem to partially confirm Kyle and Roll’s (1986) proposition that prices during trading hours, when the market is more active, are more volatile than in non-trading hours. However, the second group of figures clearly shows that return variance increases around the lunch break, which coincides with the lowest level of market activity. This raises the importance of liquidity and suggests that, when trading intensity is “normal” for the market, its contribution to the variance is minimal. Both very active and inactive stages, however, increase the price fluctuations, at least in a rather illiquid market, such as the EU ETS.

The Impact of Trading Intensity on Spread and Variance components

The next two groups of figures (6.4, 6.12, 6.20 and 6.28 is the first group, while 6.5, 6.13, 6.21 and 6.29 is the second) provide a deeper insight into the relation between spreads and return variance and levels of price volatility, informed trading and number of OTC transactions across different levels of expected trading intensity. In more detail, panel A of these presents the variations of spreads and return variance across expected trading intensity, $E[s_t|H_{t-1}]$, as in Eq. (6.12), according to the estimates in the last column of Table 6.1, and the expected level of informed trading, $E[\tilde{I}_t|H_{t-1}]$, as in Eq. (6.14). Next, panel B presents the variations of spreads and return variance across expected trading intensity and expected number of OTC transactions, which is computed as a moving average of the OTC transactions during the last fifteen minutes. Similarly panel C focuses on expected trading intensity and the expected price volatility, $E\left[\sigma_{p,t}^2|H_{t-1}\right]$, as in Eq. (6.13).

The main finding in these graphs is that the main determinant of spreads and return variance variations is the expectations concerning trading intensity, while the other variables fail to exhibit significant influence. More specifically, in all graphs, when dealers expect the market to be in an inactive stage, spreads and, consequently, return variance are relatively high. When the market is expected to be in the “fundamental” regime, however, they decrease to their minima, while as the expected trading intensity increases they increase as well, reaching their maxima when the market is really active. In parallel, spreads and return variance increase, independently of the level of expected trading intensity, when market participants expect more informed traders, more OTC traders and higher price volatility.

The effect of these variations, although rather moderate compared to trading intensity, should not be ignored. Spreads and return variance increase even further when past transactions result in expectations that constitute a riskier market environment. Increased risk can be due to higher price volatility or greater presence of informed traders. The same happens when an increased presence of OTC transactions is expected. The first two factors have been found, in the last column of Table 6.1, to affect mainly the transitory component, φ_t, where the informational component, θ_t, is mainly determined by trading intensity. OTC transactions, however, seem to have a considerable impact on theta, indicating a potentially increased informational content of these trades. This confirms the findings in the previous Chapters that connect registering of OTC positions with unresolved price information. The effect of OTC transactions becomes stronger when trading intensity expectations are higher as well. In other words, when the market environment is volatile and the market is expected to be in a high trading intensity stage, both spread and variance reach their highest levels. This raises even further the importance of managing both size and time of trading.

The next two groups of figures (the first group consists of 6.6, 6.14, 6.22 and 6.30, while the second consists of 6.7,6.15, 6.23 and 6.31), try to identify the source of, the afore-mentioned, spread and return variance variations by focusing on the impact of expectations on the adverse-selection, θ_t as in Eq. (6.16), and the trading cost, φ_t as in Eq. (6.18), components. Both groups are organized as the previous two groups: panel A presents the variations of Theta and Phi across expected trading intensity and expected level of informed trading; panel B focuses on their variations across expected trading intensity and expected number of OTC transactions; panel C examines the importance of expected price volatility.

Similar to previous findings, expected trading intensity is the main determinant of both components. Theta appears to be an increasing function of expected trading intensity, where the rate of change is different in each regime. In contrast, Phi is found to be relatively high when the expected trading intensity is in the uninformed regime. It drops sharply in the “fundamental” regime, increases sharply in the “informed” regime and then decreases moderately as expected trading intensity increases.[185]

Combining all the above information, spreads and return variance fluctuate because of the relative combination of theta and phi. The main determinant of these two components appears to be the expected trading intensity, since the expectations of the other variables fail to provide radical changes. Consequently, theta and phi seem to summarize the informational, and thus permanent, and the liquidity, and thus transitory, content of trades respectively. Therefore, when the market is expected to be in an

[185] Please note that the thresholds for identifying these regimes are derived from the analysis in Chapter 5.

inactive stage, uninformed, the information component is low, but the cost of holding inventories (or portfolios) is high, because it is more difficult to adjust positions to the optimal level. But when the market is expected to be in a mid-stage of activity, "fundamental" regime, spread and return variance reach their lowest levels. The main reason for this is a very low liquidity cost component, $\varphi_1 + \sum_{\pi}^{3}\left(\varphi_{2,\pi} I_{\pi,t}\right) E[s_t|H_{t-1}]$, and thus a small transaction cost, φ_t. In the "fundamental" regime this component is at its lowest level and counteracts the relatively small increase in the information component, θ_t. If expectations belong to the highest regime, informed, spread and return variance increase considerably. The main source of this increase is theta, which surpasses the relatively moderate diminishing effect of increased liquidity in the transitory component. The further the transaction is expected to be from the threshold, s_2 as in Eqs. (5.7) and (5.8), the greater is the informational over the liquidity impact on prices.

6.3.3 Type of order

Limit versus Market orders

Another important point that has been mentioned in Section 6.1, is the choice of the type of order to be submitted. Considering that investors should submit market orders when they expect the "true" price to change (i.e., when there is unresolved information that has yet to be incorporated into prices), and limit orders when they expect moderate "true" price changes (where they can earn the spread from liquidity traders), trading intensity could potentially help in determining the right strategy. In more detail, trading intensity has been found to play a dual contradictive role, which is reflected in theta and phi. After a transaction, dealers formulate expectations about trading intensity and risk. These expectations revise their beliefs about the "true" value of the asset, and consequently about any unresolved price-relevant information, and about future liquidity and risk. The first alters the compensation they require for trading with better informed traders, measured by θ_t as in Eq. (6.16), while the second revises their price for supplying liquidity, measured by φ_t as in Eq. (6.18). Previous findings have shown that these changes are mainly caused by trading intensity and, therefore, that it could help in choosing the appropriate trade type. Assuming they are rational uninformed traders who observe market history trying to learn, they should be expected to submit a market order when they expect radical "true" price changes, measured by large θ_t, since the spread might not be enough to cover a high probability of an "unfavourable" price change. In contrast, when θ_t is low, price changes are expected to be moderate, and thus the spread, after submitting a limit order, should be profitable.

Looking at the variations of theta and phi across expectations, in Figures 6.4, 6.12, 6.20, 6.28, which refer to half spread variations, 6.6, 6.14, 6.22, 6.30, which refer to Theta variations and 6.7, 6.15, 6.23, 6.31, which refer to phi variations, suggests that limit orders should be preferred when the market is expected to be in the uninformed regime, while market orders should be preferred when the market is expected to be in the next two regimes. More specifically, in the uninformed regime, in quiet stages of the market, theta is very low, while phi is quite high. This, according to the assumption presented in the previous paragraph, is the ideal environment for limit order submission. Investors can earn the spread, without bearing substantial information-related risk. In the fundamental trading regime, however, spread reaches the lowest point, mainly because of phi, and therefore, market orders become more profitable, since the informational content of trades, θ_t, increases. Along the same lines, in the informed regime, limit orders should not be expected to be profitable due to a very high theta.

The last group of figures (6.8, 6.16, 6.24 and 6.32) presents the difference in the rate of change of variance and the rate of change of spread due to variations in expected trading intensity, as in Eq. (6.31), in all markets and phases. When a transaction is expected to increase the probability of price change, and thus the variance, $Var(\Delta p_t)$, more than spread is expected to vary, the variance change, $\frac{\partial[Var(\Delta p_t)]}{\partial E[s_t|H_{t-1}]}$, is larger than spread variation, $\frac{\partial\left[S_t^{Implied}\right]}{\partial E[s_t|H_{t-1}]}$, which, according to Eq. (6.31) means that $R_t > 0$. This is when market orders should be preferred because the “extra” profits expected by the subsequent price change are expected to be higher than the spread. But when $R_t > 0$, the subsequent price changes are expected to be moderate and, therefore, spreads should be expected to be more profitable. In accordance with previous findings, these figures show that R_t is negative in the uninformed regime, which represents the majority of transactions and is associated with large spreads. This is where limit orders should be preferred. In the other two regimes, however, R_t is positive and market orders appear to be more beneficial. This finding is very important, considering that the majority of studies dealing with this issue (see, inter alia, O’Hara and Oldfield, 1986; Chakravaty and Holden, 1995; Brown and Zhang, 1997; Parlour, 1998; Foucault, 1999; Bouchaud et al., 2004, 2006; Wyart et al., 2008) finds that limit orders are associated with wider spreads. This study provides evidence that this holds only when low levels of trading intensity are expected. In the opposite case, a large θ_t would indicate a high probability of informed trading, and thus a large, probably "adverse", subsequent price change, which would deter market participants from submitting a limit order and expecting to earn the spread.

Buy versus Sell orders

Finally, Table 6.4 provides an insight into the relation between trade initiation and market variables, such as spread, return variance and asymmetric information. The table is divided into two sections, A and B. Panel A presents the average implied spread, as in Eq. 6.24, and the adverse-selection component, θ_t, as in Eq. (6.16), across different levels of transaction size and trading frequency, in all markets and phases. The threshold values for both variables have been arbitrarily chosen to indicate low, middle and high trading activity and they are the same for both markets in both phases. Furthermore, the table differentiates the relative sizes of spreads between Buyer (B) and Seller (O) initiated transactions. Focusing on each market, the first two columns present the width of the implied spread and the size of the adverse-selection component after a trade that belongs to one of the pre-described different regimes, namely uninformed, fundamental and informed. The next two columns summarize the values of spreads and adverse-selection across different levels of trading intensity of the current transaction. Panel B presents the relative values for the second moment of price change.

This allows for the examination of various issues. First, Aitken and Frino (1996) present evidence showing that Buy orders are associated with larger transactions costs, and thus wider spreads. Their proposition is confirmed by the results in Table 6.4 for both size and time in the Carbon market. When the previous or the current transaction is larger or has occurred after a short duration, the associated implied spread is higher for Buys than for Sells. This is consistent across all sections of Table 6.4. The information component, Adverse-selection, is higher as well. This indicates that, on average, Buyer initiated traders are more likely to be driven by information, or at least dealers, or limit order traders, seem to believe so and require a higher compensation in the form of higher θ_t. However, panel B suggests that variance is consistently larger after larger Sells, which could probably be attributed to a larger interaction component. This could in practice mean that when good news hits the market, informed traders initiate Buy orders before everyone else. Their actions are observed by dealers who start incorporating the "extra" information into prices, and who buy, increasing the adverse-selection component, θ_t.

Consequently, spreads are larger. However, when bad news hits the market, informed traders sell, or short sell, and dealers once again increase the information component according to their expectations, but this time their revision is more conservative. This can be attributed to the “Buy” character of the market (Figure 5.5), which can create “false optimism”. Companies will always need to buy emission allowances, so there is a strong rather inelastic demand.[186] However, the high variance after large Sells indicates that not all traders react in the same way after large Sells. Some might under-react (Kaswan, 2010) charging a lower spread, or some might over-react by changing their quoted prices significantly.

Along the same lines, Hendall et al. (1997) postulate that large Buys convey more information than large Sells. Similarly, Chan and Lakonishok (1995) argue that large block trades have a higher informational impact. In both cases, a larger informational content of previous trades should result in a higher adverse-selection component, theta. Indeed, evidence presented in Table 6.4 suggests that when the previous transaction is either larger or occurred after a short period of time (i.e., the intensity of trading is high) theta is consistently higher, especially when trades, current or previous, are Buyer initiated. In contrast, Grammig et al. (2007) argue that after an inactive stage of the market, the next transactions are more informative, therefore higher thetas should be observed. However, a consistently larger asymmetric information component after higher trading intensity, observed in both panels of Table 6.4, is in opposition to this and tends to support the proposition of Chan and Lakonishok (1995), who connect higher demand for immediacy, and thus higher trading intensity, with larger informational impact and, therefore a larger adverse-selection component, θ_t.

Summarizing, the findings above suggest that large, high frequency transactions indeed carry more information and have a higher permanent impact on prices, and thus wider spreads are observed. Furthermore, considering the direction of trade, Buys could be associated with good news and Sells with bad. Therefore, when there is a positive information shock, traders seem to act quickly and in large volumes, buying the asset. Their actions though are observed by other market participants, such as dealers, who now require higher compensation, due to a higher asymmetric information component, and therefore the spread widens. In the opposite case, when bad (good) news hits the

[186] Both positive and negative information shocks are assumed to be equally strong. A further analysis could examine the degree of the information shocks and their different price impacts. However, that would be far beyond the limits of this analysis. For a detailed presentation of relevant studies please refer to Vives (2008).

market, not all traders appear to act the same way and therefore higher (narrower) spreads, but also higher variance, are observed. Finally, considering that OTC trades are relatively much larger than the normal trades (Figure 5.4.B), they can be easily associated with these findings, concerning large Buys or Sells, and their increasing impact on duration, trading intensity, spreads and return variance price can be justified.

6.4 Summary

The analytical focus of this study emphasizes the price impact of trading intensity, as it is measured by both transaction size and trading frequency. Different dynamics of these two variables, such as different regimes and their permanent or transitory impact, extensively discussed in the literature, are taken into account. Since the early microstructure literature, volume of trading has been utilized to proxy various market characteristics, such as liquidity, the informational content of a trade and the probability of informed trading. It has gained particular attention in the inventory models as a determinant of prices and spreads. More recently, the role of time and the informational content of inter-trade durations have gained particular attention. Especially the role of time in information dissemination and learning has been extensively studied, both empirically and theoretically. Moreover, many theoretical models have examined the dynamics of learning and formulating expectations, as well as how they affect price changes and therefore volatility, taking into account investors' risk aversion. Furthermore, other issues, such as the relation between particular market characteristics and spread/variance or the preferred order type, have also been examined.

This study extends Madhavan et al.'s (2007) idea towards Angelidis and Benos (2009) and Ben Sita (2010), proposing a new dynamic expectations joint model. The intraday price formation process is assumed to be determined by a permanent component, related to information and a transitory component, related to liquidity and risk aversion. Trading intensity plays a dual role that is allowed to have a different impact on price components. First it is used to proxy the informational content of trades and second, it acts as a natural measure of liquidity. Furthermore, prices, and consequently spreads, are recognized as being determined by expectations, concerning risk, information and liquidity. Trading intensity is used to proxy the informational and the liquidity content of trades, while the expected level of informed trading and the expected price volatility, as well as how long dealers can expect to be exposed to it, are used to measure risk. Consequently, the formulation of the expectations is of particular importance.

Expectations of informed trading and price volatility are computed as moving averages for the last fifteen minutes, while expected duration is calculated using the STM-ACD model discussed in Chapter 5. Concerning trading intensity, a linear autoregressive function is assumed to summarize its dynamics and is modelled jointly with return. This framework could be extended to include more parameters or potential non-linearities. In this way, expectations, and consequently price components, are revised after every transaction, which allows price-related variables, such as spread and variance, to be modelled as continuous variables.

In conclusion, the formulated expectation of trading intensity is found to be an important determinant of intraday price formation, especially when it is expected to be in the fundamental or the informed regimes. The theoretical predictions of its dual role, are also confirmed, since it affects in a different way the permanent and the transitory pricing components. The permanent price component is found to be an increasing function solely of trading intensity, where different regimes of trading intensity change the sensitivity of price change. The transitory price component, however, can be decomposed into a constant order-processing component, a decreasing function of trading intensity that accounts for liquidity considerations and a risk aversion component measured by the expected price volatility levels and the time dealers expect to be exposed. This transitory component reaches its lowest point in the fundamental regime and its highest point in the informed regime. Both components, which are updated after every transaction, determine spread and return variance, which are found to exhibit the usual U-shape intraday pattern found in other markets. All these relations seem to strengthen when price volatility, the level of informed trading and the presence of OTC transactions are higher. Moreover, surprisingly enough, although trading intensity appears to have a significant impact on prices, there is no sign of endogeneity, since return is not found to have a significant impact on trading intensity. In addition, the common conviction that limit orders are associated with and preferred when there are wide spreads, is confirmed, but only for the uninformed regime. In the other two regimes, namely fundamental and informed, market orders are more appropriate due to a high adverse-selection component that indicates that it should be more profitable to follow “true” price changes than trying to earn the spread. Finally, large (i.e., high transaction size), high frequency (i.e., short duration) transactions seem to carry more information and have a higher permanent impact on prices. Also, Buyer initiated transactions are related to wider spreads, while Seller initiated trades are associated with narrower spreads and higher variance.

Chapter 7

Summary, Limitations and Future Research

7.Summary, Limitations and Future Research

7.1 Non Technical Summary

Since UHF data has been available to both practitioners and researchers, financial theory has been seen from a whole new perspective. In particular, the intriguing area of price formation has shifted from a macroeconomic perspective to a transaction level focus, creating a new field of study that examines the micro-foundations of intraday trading activity. Market microstructure, as it is called, focuses on the determinants of intraday, microstructure, phenomena, employing all available public information of the highest precision, since every transaction is now recorded. On this level, behavioural aspects of market participants are taken into account, allowing the desires/needs of investors to have an impact on various market stylized facts.

More precisely, one branch of literature approaches intraday price formation from market makers' perspective, underlying the importance of their inventories. Market makers are recognized to face continuous exposure to incoming order imbalances. Orders of any direction might outnumber the opposite orders, creating an imbalance that needs to be covered by the dealer's portfolio. The risk of exposure is twofold, either in the form of carrying cost, in case of excessive inventory, or in the form of loss of sales, when dealers' inventories are not sufficient to support sales. Considering that market makers need to be ready to transact (i.e., immediacy), they need to have an optimal portfolio composition to match incoming order flow, in order either to avoid market failure, or to make profit, or simply because of a competitive environment. However, when transactions make them deviate from it, they can use their price quoting to attract or deter other market participants from trading. Therefore, Bid-Ask spreads are seen as a natural outcome of trading activity, since they derive from market makers' need to compensate for providing immediacy. As soon as they return to their optimal level, prices return to their "fully" informational level.

This class of models is called inventory-based models and, although it is a rather incomplete approach, it gives a further insight into price formation. First, these models investigate further the price impact of liquidity and distinguish that particular cost component from transaction costs that are mainly fixed. The liquidity component of the spread fluctuates along the various levels of liquidity, according to the positions of dealers. Second, they recognize that market makers observe the market, through past trading activity, in order to formulate expectations concerning any incoming order flow imbalances. Consequently, they see the intensity of trading as an undeniable market determinant. Third, they implicitly assume that there is a time dimension of the arrival of trades, by relaxing the assumption of fixed time intervals. The time of occurrence of transactions is studied separately and it is related to price-relevant information. Although the probability of the direction of a trade might be 1/2 over longer periods of time, excessive imbalances might occur and order flow might be highly autocorrelated on shorter periods over the trading day. Fourth, inventories are described as a "buffer" to market activity, since these price revisions are driven only by inventory needs, which are unrelated to information. Price variations occur in case of persistent order flow fluctuations and the quotes return to their "fair" values after the end of the episode. Consequently, the price impact of liquidity is understood as transitory, where competition, in the form of higher number of market makers or the type of orders (limit orders are seen as market making in one side of the spread), might shorten that effect even further.

Another branch of literature emphasizes the informational content of trades. These models maintain that spreads would exist even without explicit trading costs due to asymmetric information. They argue that market makers will always lose money when dealing with better informed traders. Their only compensations can be in the form of earning the spread from trading with uninformed traders. Considering that they cannot be in a position to know the level of private information in the market, they can only speculate. Therefore, they observe market activity, in the form of past trading, they learn (PIN is a key concept that summarizes their beliefs) and they formulate expectations concerning the level and the content of price unresolved information. This constitutes the information component of the spread, which is seen as their compensation for their "immediacy" and "lack of knowledge". This, along with order-processing and inventory-holding, is the third most widely recognized component of the trading spread.

The contribution of information-based models, as they are referred to in the literature, is multidimensional. First, they recognize that there are different types of traders according to their access to price-relevant information and that their trading behaviour has a significant price impact. According to these models, informed traders possess information about the "fair" value of the underlying asset, prior to everyone else. They have the incentive to exploit the inforamtion against uninformed traders. Their choice to act immediately or strategically depends on the market setting and the level of competition. Furthermore, uninformed traders are divided into two categories. “Discretionary”-liquidity traders observe the market and try to learn from past trading history, while “non-discretionary”-liquidity traders possess only public, price resolved information and they trade for reasons other than the arrival of exogenous information. Consequently, intraday price formation is understood as an equilibrium of the proportion of these market participants, considering the strength of the information signal. Second, similar to inventory models, there is an implicit consideration of time. Informed traders possess their information before everyone else and they reveal it progressively through their trading. Therefore, there is a period of time from the information arrival until it is fully resolved into prices.

Third, this space in time allows other traders to learn, but their learning might vary both in terms of quality and speed, although they observe the same events. This point raises significant issues. The event they observe to formulate their expectations about PIN refers to information that can be extracted by past trades. More precisely, they recognize that there are actions that reveal future price changes, due to information that is not incorporated into prices yet. This way these models implicitly assume that there is price-relevant information in non-price variables, such as order flow and the intensity of trading. Consequently, trading history is available to all market participants, but not all traders interpret it the same way. When aggregating information, they might have different portfolio needs, either because of different quality or different quantity thresholds, or simply different learning speeds. In addition, the informational, and thus permanent, content of trades cannot easily be distinguished from the liquidity impact, which is transitory. According to the last point, these models also recognize that price revisions due to information alter the beliefs of market participants concerning the “fair” value of the asset and therefore it is permanent.

Although inventory and information models offer a better understanding of market microstructure, they both operate under a very strong limitation. Their approach is fairly theoretical and they understand the market as a “game” with different players. However, the “prior knowledge” of the game or the “set-up” of the framework, describing market structure, market participants and other similar issues, must be known. Obviously, that framework is not appropriate when empirical issues are of interest. Therefore, another class of models has emerged that aims at explaining price changes, and consequently the determinants of Bid-Ask spreads, by employing a stationary covariance assumption.

These models have become very popular due to their simplicity of estimation, and due to the fact that spreads can be derived directly from transaction prices. Furthermore, another source of uncertainty is "re"-introduced to account for order flow variations, introducing a new class of models; the trade indicator models. These models try to derive spread components by assessing the direction of trade.

In parallel, several other, empirical, aspects have been discussed in various empirical models. One of the most important issues is the informational or liquidity content of trades and how that can be extracted by past trading. The literature emphasizes the role of trading volume and trading frequency. The trading behaviour of informed traders is related to the arrival of information. When there is no price-relevant information they do not have an incentive to transact and therefore when they enter the market they increase the volume of trading. Consequently, increased trading volumes are related to informed trading. Along the same lines, higher trading frequency could indicate the same thing. Time has also been connected to the quality of information, and in particular to the arrival of good or bad news. More recently, several studies emphasize empirically the price impact of trades and they try to derive permanent and transitory price components by either transaction size or trading frequency.

Time has gained considerable attention. Unlike data sets of lower frequency, when transaction level data are employed, time is irregularly spaced. Several studies underline the importance of its modelling, because it violates the basic assumption of fixed time intervals, present in empirical models, and because it might convey price-relevant information. Furthermore, time series of duration are persistent and over-dispersed. Considering these stylized facts of time, Autoregressive Conditional Duration (ACD) models have been developed. They model duration as a dependent point process. Modelling needs a conditional mean specification that depends on past durations and a conditional density function with a positive support.

The main objections to the initial model concern the simplicity of the specifications. Several studies maintain that non-linear specifications for the conditional mean should be more appropriate for describing the Data Generation Process (DGP) of durations. Along the same lines, distributions of higher complexity are employed in order to capture more precisely higher moments of duration series. More recently, ACD models have been connected, empirically or theoretically, to informed trading. The trading activity of different types of traders should follow different patterns that can be described by the arrival of their trades. Statistical properties of the duration series could reveal their presence.

The application of microstructure concepts, and especially whatever refers to information and price resolution, is particularly relevant in young and illiquid markets, because information is an important price determinant. This last comment is important in the Carbon market due to its unique stylized facts. It is politically influenced and therefore an appropriate regulatory framework is needed to ensure the right balance between market innovation, price informativeness and liquidity, in order to achieve its final goal, which is to make emissions expensive. However, this market has only recently gained academic attention, which emphasizes mainly regulatory and market structure issues. Furthermore, several studies analyze the nature of the asset, the price formation between spot and futures contracts, as well as the determinants of price volatility. They mainly use daily data sets. More recently, several studies examine microstructure issues, such as intraday price formation, spread decomposition and price leadership, but the literature remains rather sparse.

Several microstructure issues have not been examined in the Carbon market. It appears that there is no relevant study that models time since the inception the market. This is the primary concern of Chapter 4 of this thesis. In this section, time is explicitly modelled using various specifications of ACD models. The market differs significantly from other more liquid and well established financial markets and this might have an impact on the time the next transaction is expected to occur. Two main market features that are of particular interest are the consistently increasing trading volume, as market gains complexity and the OTC transactions. Trading volume, in terms of number of transactions, as well as aggregated volumes, has increased considerably over the last years and exhibits strong seasonalities. In addition, it is a buyer dominated market, with periods of illiquidity, especially in early stages, which magnifies potential price impact of microstructure effects. In addition, OTC allowance holders are allowed to register their positions in the organize market to mitigate counterparty risk.

The models proposed in this section contribute to the literature in various ways. First, they extend the Carbon market literature, by further investigating intraday market dynamics of the trading process. Previous studies emphasize the lead-lag relation between spot and future prices, as well as its time dependence and clustering, focusing mainly on price dynamics. This study focuses on the trading process in terms of trading patterns and models time, trying to incorporate market stylized facts in order to explain potential autoregressive dynamics of market liquidity. Second, this study contributes to the ACD literature as well, by suggesting two empirical adjustments in order to account for market unique features. This adds to the debate challenging the simplicity of the model specifications arguing that further empirical adjustments should be employed in order to better describe the Data Generation Process (DGP) of duration, especially when the market environment is rather illiquid and significantly differs from other organized markets. Finally, the proposed models explicitly incorporate other variables in duration modelling, in a way that does not challenges the time endogeneity assumption of ACD models. This extends the ACD literature by providing a framework, which allows for the inclusion of other variables, according to the particular modelling needs of the market under investigation.

In more detail, in this chapter the performance of several ACD specifications is examined, focusing on whether empirical adjustments further improve forecasting and fitting accuracy. Two new models are proposed that allow for asymmetric impact of past duration on expected duration, without challenging the assumption that time is endogenously determined. Expected duration still depends on past durations, but their effect is not constant. It might be asymmetric depending on other variables. The choice of these variables depends on the market environment and this framework allows for numerous specifications. The first model is a non-linear specification and focuses on the potential asymmetric effects of past durations. The size of the impact of past realized arrival times is allowed to vary across different regimes of an economic relevant variable. In this study trading intensity is chosen.[187] The second model allows the impact of past durations to be adjusted on a piece-wise linear fashion, depending on the regimes of an economic relevant variable. The overall impact is asymmetric and non linear, but the regime-specific relation is still linear. In this study a binary (dummy) variable is employed to account for the presence of OTC transactions.

The main finding in this chapter indicates that, although ACD models perform well in this new market, empirical adjustments significantly improve model performance. This is also consistent with the proposition of Bauwens et al. (2004), that the conditional mean specification contributes more to fitting and forecasting than the distributional assumptions. In more detail, more complex density functions better describe the DGP of

[187] To proxy trading intensity, in accordance with Engle and Russell (1998) and De Luca and Zuccolotto (2006), a natural measure is employed, which is the ratio of trading size to duration. This measures the number of contracts traded per unit of time.

duration, but the main improvement comes from the conditional mean specification. Precision of the models is further improved after accounting for potential asymmetries. Trading in the most liquid ECX appears to be influenced by innovations in trading intensity, while OTC transactions have a more profound effect in NP. In addition, although expected duration seems to asymmetrically depend on past durations, the linear approach fits and forecasts better, especially on the long term. Both parameterizations capture the illiquid shocks in the market. The non-linear parameterization suggests that there is a market momentum. Large and frequent transactions increase market liquidity and the expected duration decreases. However, OTC transactions appear to increase the expected duration, either because they carry information and deter other traders from transacting, or because they consume the current levels of liquidity. The market might then need some time to reach a new equilibrium. These two combined provide an initial indication of a market that is sensitive to liquidity and information.

These finding are particularly relevant to the trading process in the European Carbon market and they could improve the practices of the following groups. First, a better understanding of intraday trading activity can enhance market regulation in reaching a balance between market innovation and liquidity needs. According to Viswanathan (2010), this would provide the foundations of more accurate pricing and would help the market to serve its purpose of reducing emissions. The market appears to be sensitive in both information and liquidity and an accurate, empirically adjusted, duration model could provide a natural measure of sensitivity towards both. This can improve monitoring and consequently effectiveness of regulation. In addition, a more precise duration model could also improve market making practices in EU ETS. Market makers can better manage risk if they know how long (i.e., expected duration) they are going to be exposed to it. This could contribute further to market efficiency, since it would result in narrower spreads. This applies to limit order traders, who can also benefit from a better duration model. Their trading resembles market making on one side of the spread, and thus they can develop better trading strategies by managing the time dimension of “execution” risk.

In Chapter 5, the trading patterns in the European Carbon market are further investigated, relating duration with information and volume. In more details focus shifts to the informational content of trading intensity, and the role of duration modelling. A new ACD framework is proposed focusing on connecting the probability of a transaction to occur now, given that it has not until now (i.e., hazard function) with different types of trades and, consequently, different types of traders. A key idea is the role of time in information resolution, and more precisely when and for how long a piece of information is exploitable. Unlike previous literature, the proposed model focuses on the process of information resolution rather than on its content. The quality, the strength and the direction of information are not modelled explicitly, but they are assumed to be progressively revealed by traders' actions. Market is seen as being efficient, in the sense that rationally incorporates information, but imperfect, in the sense that it takes some time from the moment that information hits the market up to when prices reflect that information. This period of "price adjustment" provides the opportunity to some traders to act before everyone else and make a profit. The prices do not fully incorporate available information, even if it is public, and in fact given market conditions, this period might be prolonged. However, other observable variables, such as trading intensity might convey that information. Consequently, the information benefit is now translated into a "timing" benefit. Modelling such time dimension of information allows for identifying informed trades/traders in a natural way, derived from past observable information. The "time dimension" of information along with the proposed framework that can attach a probability to a transaction to be informed constitute the main contributions to the literature.

Drawing on De Luca and Zuccolotto (2006) and Hujer and Vuletic (2007), a Smooth-Transition-Mixture of Distributions ACD (STM-ACD) model of duration is proposed. Teräsvira's smooth transition framework is applied to the shape parameter of a Weibull distribution, allowing it to vary across three different regimes of an economically relevant variable. This determines the shape of the distribution (the Weibull nests the Exponential) and consequently the shape of the hazard function. The three recognized regimes correspond to three different types of trades/traders; uninformed, fundamental and informed. The smooth transition function accounts for the learning process between regimes and therefore for hybrid trades. Following Hujer and Vuletic (2007), the shape of the hazard function is then related to a particular type of trades/traders, according to its slope. However, unlike previous work, trading intensity, which is observable, is used as threshold variable, and this allows the identification of the regime to which each transaction belongs. Further, focusing on informed trading activity, the probability of the next transaction to be informed can be computed after forecasting the level of future trading intensity.

The most important finding of Chapter 5 involves the role of trading intensity in transmitting information signals. First, trading intensity appears to be sufficient in identifying three distinct regimes. Then, according to the shape of the hazard function, the STM-ACD confirms the theoretical propositions of Easley and O'Hara (1992) and Dufour and Engle (2000b) that increased trading activity is associated with information. According to the shape of the hazard function, low trading intensity, in the form of either low trading size or long duration, is related to uninformed trades/traders, while high trading intensity is related to more informed trading. The higher the trading intensity, the more informed the trade. Consequently, large or fast transactions are interpreted as closely related to information and are, therefore, expected to have a greater price impact. Similarly, based on the findings of Pascual et al. (2004), trading intensity increases following informed trades, but this happens because of higher trading frequency rather than higher trading size. In addition, empirical findings provide further evidence that longer durations are associated with no news, according to Easley and O'Hara's (1992) propositions, and contrary to Diamond and Verrecchia (1987), who argue that longer durations are associated with absence of new information. Findings also support Kyle (1985) indicating that informed traders act strategically by segmenting their trades.

The discussion above is particularly important and relevant to various aspects of trading practice in the European Carbon market. First, it highlights the benefit of possessing information before it is fully incorporated into prices, recognising that there are non-pricing variables that can reveal information. On microstructure level, market imperfections are exploitable, even for a very short period of time, and that increases the intrinsic value of acquiring information in "real time". This can affect the attitude of market participants towards the timing and the cost of acquiring information. Second, an increased ability of market participants to indentify informed trades by simply observing past transactions, can increase their market power. Given the "buyer" character of the market, they should easily compensate for potential losses by trading with better informed traders. Their main concern should be liquidity, and therefore spreads should only increase when they observe increased information-based trading and insufficient liquidity. On overall, lower spreads should be observed, and this contributes to increased market efficiency. Third, the proposed model can be used for monitoring purposes by regulatory authorities. By identifying informed trading, further action can be taken to protect the market from "manipulation". This can also be applied in real time to adjust the balance between market innovation and liquidity.

In Chapter 6 the focus shifts on the application of the previous findings. Market makers and limit order traders formulate expectations and they quote their prices accordingly. Using the models discussed in the previous two chapters they can forecast expected duration and they can attribute a probability of the next transaction to be informed by forecasting the level of activity the market is going to be in. Therefore they can have an estimated time that they will be exposed to various type of risks, as well as they can anticipate the price impact of the incoming trade. A key aspect of this process is the dual pricing impact of trading intensity. According to previous models duration and transaction size have an undeniable impact on prices, both permanent and transitory. In this study, their post-trade price impact is further investigated, by allowing expectations to affect the information and liquidity component of price change. This can be interpreted as a natural measure of market sensitivity towards information and liquidity, which has been discussed in the previous chapters, and could explain the conflicting results in the literature.

In more detail, a new dynamic expectations structural model for intraday price changes is proposed, extending the existing literature in the following ways. First, it deviates from a static approach of modelling intraday prices and Bid-Ask spread, in the sense that it allows price components to vary according to trading activity. The information and liquidity components are revised after trading intensity fluctuations. Second, price revisions depend on the, Bayesian, learning process of market participants, who condition their quoted prices upon expectations about the price impact of future transactions. Consequently, price and spread components are modelled as continuous latent variables that depend on past trading history, public information and market participants' learning ability. Third, the dynamic character of the model provides a framework to measure the information and liquidity pricing impact of a trade, in a way that the profitability of a market or limit order strategy can be maximized. These components are measurable and they are revised after every transaction. Investors can formulate expectations and then take an appropriate market position. Finally, this model investigates further the intraday price formation of Carbon allowances, recognising the pricing impact of market participants' behaviour. This can have numerous implications to regulation and trading in EU ETS.

The model proposed in this chapter extends the original model of Madhavan et al. (1997) and draws on the studies of Grammig et al. (2007), Angelidis and Benos (2009) and Ben Sita (2010). Market makers, in the context of a hybrid market, continuously observe trading history through trading intensity, trying to extract price-relevant information. They quote their prices based on liquidity considerations and the post-trade effect of their actions. Similar to Madhavan et al. (2007), intraday returns are driven by both an information and a liquidity component. In this thesis, however, these are allowed to be determined by dynamic expectations of trading intensity, and by the expected exposure to risk. The employed risk measures account for both level of risk and time of exposure. Thus, the model can be used to define the components of estimated spread and intraday volatility, as well as to help identify the most appropriate type of order, between Market and Limit Orders.

One of the main findings is that spread components follow intraday patterns that seem to be adequately explained by the dual role of trading intensity. Trading intensity is positively related to the information component, and, probably due to economies of scale or the illiquid character of the market, expectations of higher trading intensity have a decreasing effect on price changes. The higher the expectation, the higher the price impact of a trade, which confirms the empirical findings of Dufour and Engle (2000b). In addition, and consistent with the literature, risk, in the form of price volatility, seems to affect only the liquidity component of the spread and not the information component. Another, main finding is the confirmation of the positive relation between limit orders and spread width. However, it appears that a limit order strategy could be profitable only when trading intensity is low. Otherwise, the information component indicates that the midpoint deviation would be larger than the spread. Finally, Buy orders, especially the large ones, are associated with a higher adverse-selection component and wider spreads than Sell orders.

These findings are particularly relevant to various aspects of trading process in EU ETS. First, trading practices can be improved, as investors could measure more precisely the price impact of their trades, which might vary over time. They can account for behavioural aspects of trading in real time and they can formulate trading strategies that take into account the actions of other investors. These actions can be included in the expectations' equation and thus can be explicitly modelled. This provides a promising framework that connects the theoretical approach of "inventory" and "information" models with the empirical flexibility of “trade indicator” models and the forecasting ability of “VAR” models.

Further, this practice can also be beneficial to investors who possess price unresolved information, as they can adjust their trading to current liquidity levels and sensitivity towards new information, and maximize their profits. These traders can measure the post-trade price impact of their actions and they can adjust their strategies accordingly in order to minimize their visibility and maximize their profit. Second, market making can be further improved, since this model can indicate when market orders can be profitable. Dealers can adjust spread width according to liquidity and information and they can become more efficient in managing both. This would result in narrower spreads and further contribute to market development and maturity. Finally, a better understanding of trading practices and their pricing impact can improve regulation and monitoring, and thus the market can be more efficient in achieving its goal, to reduce emissions. This model proposes a natural measure of market sensitivity towards information and liquidity, which are both highlighted in the Carbon market literature. Regulatory authorities can develop policies that can manage both and allow the market to reach a balance between “market innovation” and liquidity.

7.2 Limitations and Future Research

The empirical findings presented in the previous three chapters, although they support the validity of the models proposed, are recognized to be extracted under some quite restrictive assumptions. In the first empirical chapter (Chapter 4) duration is modelled solely as an exogenous variable. This is in line with Engle and Russell (1998) and Engle (2000), but several studies, even in Dufour and Engle (2000a, 2000b), raise concerns about the validity of this assumption. Expected duration is only conditioned on past arrival times and no external variable is allowed to have a direct impact on its DGP. Trading intensity is highly correlated to duration, since it is derived from it, and it affects only indirectly the magnitude of the coefficient that measures the size of the impact of past durations. The variable that accounts for OTC transactions is a dummy variable and it affects only the magnitude of the coefficient that measures the impact of past durations. Furthermore, contrary to Bauwens et al. (2004), who argue that simple ACD models cannot capture duration’s higher moments, a single distribution is considered to be sufficient in modelling duration series. In addition, although several distributions of increased flexibility are examined, they all belong to the same family, and they all require the same restrictions.

In the second empirical chapter (Chapter 5), the chosen distribution, although it helps identifying differently shaped hazard functions, belongs to the exponential family and it can only allow for specific shapes. The Weibull distribution appears to be a flexible and convenient framework for the analysis that follows, but it is recognized to produce only monotonic hazard functions. In addition, the choice of an observable variable might be too simplistic for generalizing. It might be sufficient for empirically examining the informational impact of trading intensity, but by no means, can proxy all relevant information. Moreover, the assumption that only the last trade has an impact on expected duration, while any informational content of previous transactions is ignored, is rather strong. The proposed model is an empirical investigation of the informational content of trading intensity and it is a convenient compromise in the trade-off between flexibility and applicability. It cannot summarize all available information and therefore it suffers from loss of generality. But trading intensity is observable and its future values, and thus the regime the next transaction is expected to be, can be forecasted.

The main restriction in the model discussed in the third empirical chapter (Chapter 6) is the assumption that only price-relevant information is included in the last transaction. Asymmetric or asynchronous effects of past trading are assumed to be summarized on the last, post-trade, efficient price, which is unobservable. According to Hasbrouck (2007), depending on the context, this might be appropriate or too restrictive, but in any case it is a structural approach, mainly aiming at describing the trading process. Along the same lines, although the proposed model tries to connect the VAR framework with a structural model, through formulating expectations that depend on past trading history, it is still a structural approach. The dynamic character of the trading intensity expectation, allows a continuously updating process, but the only level of price-relevant information recognized is summarized in the last transaction. Moreover, for simplicity of the estimation process, the other two expectations, namely duration and price volatility, are assumed to be completely exogenous. This could be a rather strong assumption, especially when liquidity and risk are of great importance.

However, the limitations of this study can work as a facilitator for future extensions of that framework. Some ideas that could be further developed, first refer to the model proposed in the second empirical chapter. The choice of the threshold variable might be too restrictive. A latent variable, although it would increase the complexity of the estimation process, would be more appropriate in generalizing the empirical findings. The additional source of uncertainty could potentially be jointly modelled with the density function of duration using copulas. Other variables could also be included as determinants of potentially different regimes of duration, generalizing both in time and in the informational content of other marks. In addition, a distribution with higher flexibility, such as the Generalized Gamma or the Burr distribution, could provide non-monotonic hazard functions that allow for deeper interpretation. The additional free parameters could be allowed to be affected by different variables, creating a multi-dimensional modelling. The model in the third empirical chapter could be extended towards a more dynamic approach, in which all expectations can be modelled endogenously and thus jointly. This would allow for a better description of the trading process. Further, the existing framework could be extended to take into account longer memories, in order to enforce the link with the time series (VAR) approach.

Bibliography

ADMATI, A. R. & PFLEIDERER, P. (1991) Sunshine trading and financial market equilibrium. Soc Financial Studies.

ADMATI, A. R. (1985) A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica: Journal of the Econometric Society,* 53**,** 629-657.

ADMATI, A. R. & PFLEIDERER, P. (1988) Selling and trading on information in financial markets. *The American Economic Review,* 78**,** 96-103.

ADMATI, A. R. & PFLEIDERER, P. (1988) A theory of intraday patterns: Volume and price variability. *Review of Financial Studies,* 1**,** 3.

ADMATI, A. R. & PFLEIDERER, P. (1989) Divide and conquer: A theory of intraday and day-of-the-week mean effects. *Review of Financial Studies,* 2**,** 189.

AHN, H. J., CAI, J., HAMAO, Y. & HO, R. Y. K. (2002) The components of the bid-ask spread in a limit-order market: evidence from the Tokyo Stock Exchange. *Journal of Empirical finance,* 9**,** 399-430.

AITKEN, M. & FRINO, A. (1996) Execution costs associated with institutional trades on the Australian Stock Exchange* 1. *Pacific-Basin Finance Journal,* 4**,** 45-58.

AKERLOF, G. A. (1970) The market for" lemons: Quality uncertainty and the market mechanism. *The quarterly journal of economics,* 84**,** 488-500.

AKERS, R. L. (2009) *Social learning and social structure: a general theory of crime and deviance,* Transaction Publishers.

ALBEROLA, E. (2009) The EU emissions trading scheme: The effects of industrial production and CO2 emissions on carbon prices. *Economie internationale***,** 93-125.

ALBEROLA, E., CERVERO, S. G., LOPEZ, H. & UBIDE, A. (2000) Global Equilibrium Exchange Rates: Euro, Dollar," Ins,"" Outs," and Other Major Currencies in a Panel Cointegration Framework. Econometric Society.

ALBEROLA, E. & CHEVALLIER, J. (2009a) Banking and Borrowing in the EU ETS: An Econometric Appraisal of the 2005-2007 Inter-temporal Market. *International Journal of Energy, Environment and Economics, forthcoming*.

ALBEROLA, E., CHEVALLIER, J. & CHÈZE, B. (2009b) Emissions compliances and carbon prices under the EU ETS: A country specific analysis of industrial sectors. *Journal of Policy Modeling,* 31**,** 446-462.

ALBEROLA, E., CHEVALLIER, J. & CHÈZE, B. I. (2008) Price drivers and structural breaks in European carbon prices 2005-2007. *Energy Policy,* 36**,** 787-797.

ALBEROLA, R. (2007) Estimating Volatility Returns Using ARCH Models. An Empirical Case: The Spanish Energy Market. *Lecturas de Economía,* 66**,** 251-276.

ALDRIDGE, I. (2009) *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems*, John Wiley & Sons.

ALLEN, F., QIAN, J. & QIAN, M. (2005) Law, finance, and economic growth in China. *Journal of Financial Economics,* 77**,** 57-116.

ALMEIDA, A., GOODHART, C. & PAYNE, R. (1998) The effects of macroeconomic news on high frequency exchange rate behavior. *Journal of Financial and Quantitative Analysis,* 33**,** 383-408.

AMIHUD, Y. (2002) Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets,* 5**,** 31-56.

AMIHUD, Y. & MENDELSON, H. (1980) Deallership Market. Market-Making with Inventory. *Journal of Financial Economics,* 8**,** 31-53.

AMIHUD, Y. & MENDELSON, H. (1986) Asset pricing and the bid-ask spread. *Journal of Financial Economics,* 17**,** 223-249.

AMIHUD, Y. & MENDELSON, H. (1987) Trading mechanisms and stock returns: An empirical investigation. *Journal of Finance***,** 533-553.

AMIHUD, Y. & MENDELSON, H. (1988) Liquidity and asset prices: Financial management implications. *Financial Management,* 17**,** 5-15.

AMIHUD, Y. & MENDELSON, H. (1988) Liquidity, volatility and exchange automation. *Journal of Accounting, Auditing and Finance,* 3**,** 369-395.

AMIHUD, Y. & MENDELSON, H. (1991) Volatility, efficiency, and trading: Evidence from the Japanese stock market. *Journal of Finance,* 46**,** 1765-1789.

AMIHUD, Y., MENDELSON, H. & LAUTERBACH, B. (1997) Market microstructure and securities values: Evidence from the Tel Aviv Stock Exchange. *Journal of Financial Economics,* 45**,** 365-390.

AMIHUD, Y., MENDELSON, H. & WOOD, R. A. (1990) Liquidity and the 1987 stock market crash. *The Journal of Portfolio Management,* 16**,** 65-69.

AMIHUD, Y. (1990) Stock market microstructure and return volatility:: Evidence from Italy. *Journal of Banking & Finance,* 14**,** 423-440.

ANDERSEN, T. G. & BOLLERSLEV, T. (1997) Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns. *Journal of finance,* 52**,** 975-1005.

ANDERSEN, T. G. & BOLLERSLEV, T. (1997) Intraday periodicity and volatility persistence in financial markets. *Journal of empirical finance,* 4**,** 115-158.

ANDERSEN, T. G. & BOLLERSLEV, T. (1998) Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review,* 39**,** 885-905.

ANDERSEN, T. G., CHUNG, H. J. & SØRENSEN, B. E. (1999) Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics,* 91**,** 61-87.

ANGEL, D. P. (1994) *Restructuring for Innovation: the remaking of the US semiconductor industry*, Guilford.

ANGELIDIS, T. & BENOS, A. (2009) The Components of the Bid Ask Spread: the Case of the Athens Stock Exchange. *European Financial Management,* 15**,** 112-144.

ANSLEY, C. F., SPIVEY, W. A. & WROBLESKI, W. J. (1977) A class of transformations for Box-Jenkins seasonal models. *Applied Statistics,* 26**,** 173-178.

ANSLEY, C. F., SPIVEY, W. A. & WROBLESKI, W. J. (1977) On the structure of moving average processes. *Journal of Econometrics,* 6**,** 121-134.

ARTHUR, W. B., HOLLAND, J. H., LEBARON, B., PALMER, R. & TAYLER, P. (2009) Asset pricing under endogenous expectations in an artificial stock market. *Ann Arbor,* 1001.

BACK, K. (1992) Insider trading in continuous time. *Review of Financial Studies,* 5**,** 387.

BACK, K. & BARUCH, S. (2004) Information in securities markets: Kyle meets Glosten and Milgrom. *Econometrica,* 72**,** 433-465.

BACK, K., CAO, C. H. & WILLARD, G. A. (2000) Imperfect competition among informed traders. *The Journal of Finance,* 55**,** 2117-2155.

BAGEHOT, W. (2001) The only game in town. *Readings in games and information,* 21**,** 214.

BAILEY, R. E. (2005) *The economics of financial markets*, Cambridge University Press.

BAILLIE, R. T. & BOLLERSLEV, T. (1989) Common stochastic trends in a system of exchange rates. *Journal of Finance,* 44**,** 167-181.

BAILLIE, R. T. & BOLLERSLEV, T. (1991) Intra-day and inter-market volatility in foreign exchange rates. *The Review of Economic Studies,* 58**,** 565-585.

BAILLIE, R. T. & BOLLERSLEV, T. (1994) Cointegration, fractional cointegration, and exchange rate dynamics. *The Journal of Finance,* 49**,** 737-745.

BAILLIE, R. T. & BOLLERSLEV, T. (2002) The Message in Daily Exchange Rates. *Journal of Business and Economic statistics,* 20**,** 60-68.

BAILLIE, R. T., BOLLERSLEV, T. & MIKKELSEN, H. O. (1996) Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics,* 74**,** 3-30.

BAILLIE, R. T. & DEGENNARO, R. P. (1990) Stock returns and volatility. *Journal of Financial and Quantitative Analysis,* 25**,** 203-214.

BAILLIE, R. T., GEOFFREY BOOTH, G., TSE, Y. & ZABOTINA, T. (2002) Price discovery and common factor models* 1. *Journal of Financial Markets,* 5**,** 309-321.

BANERJEE, A. V. (1992) A simple model of herd behavior. *The Quarterly Journal of Economics,* 107**,** 797-817.

BARCLAY, M. J., LITZENBERGER, R. H. & WARNER, J. B. (1990) Private information, trading volume, and stock-return variances. *Review of Financial Studies,* 3**,** 233.

BARCLAY, M. J. & WARNER, J. B. (1993) Stealth trading and volatility:: Which trades move prices? *Journal of Financial Economics,* 34**,** 281-305.

BAUWENS, L. & GALLI, F. (2009) Efficient importance sampling for ML estimation of SCD models. *Computational Statistics and Data Analysis,* 53**,** 1974-1992.

BAUWENS, L., GALLI, F., GIOT, P. & UNIVERSITÉ CATHOLIQUE DE LOUVAIN . CENTER FOR OPERATIONS RESEARCH AND, E. (2003) *The moments of Log-ACD models*, Universite catholique de Louvain.

BAUWENS, L. & GIOT, P. (2000) The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks. *Annales d'Economie et de Statistique***,** 117-149.

BAUWENS, L. & GIOT, P. (2001) *Econometric modelling of stock market intraday activity*, Springer.

BAUWENS, L. & GIOT, P. (2003) Asymmetric ACD models: introducing price information in ACD models. Springer.

BAUWENS, L., GIOT, P., GRAMMIG, J. & VEREDAS, D. (2004) A comparison of financial duration models via density forecasts. *International Journal of Forecasting,* 20**,** 589-609.

BAUWENS, L., GIOT, P. & UNIVERSITÉ CATHOLIQUE DE LOUVAIN . CENTER FOR OPERATIONS RESEARCH AND, E. (1997) *The logarithmic ACD model: an application to market microstructure and NASDAQ*, Citeseer.

BAUWENS, L. & HAUTSCH, N. (2009) Modelling financial high frequency data using point processes. *Handbook of Financial Time Series***,** 953-979.

BAUWENS, L. & LUBRANO, M. (1998) Bayesian inference on GARCH models using the Gibbs sampler. *Econometrics Journal,* 1**,** 23-46.

BAUWENS, L. & VEREDAS, D. (2004) The stochastic conditional duration model: a latent variable model for the analysis of financial durations. *Journal of Econometrics,* 119**,** 381-412.

BEBCZUK, R. N. (2003) *Asymmetric information in financial markets: introduction and applications*, Cambridge University Press.

BEJA, A. & GOLDMAN, M. B. (1980) On the dynamic behavior of prices in disequilibrium. *The Journal of Finance*, 35**,** 235-248.

BEN SITA, B. (2010) Autocorrelation of the trade process: Evidence from the Helsinki Stock Exchange. *The Quarterly Review of Economics and Finance,* 50**,** 538-547.

BENNING, J. F. (2007) *Trading Stategies for Capital Markets*, McGraw-Hill.

BENSTON, G. & HAGERMAN, R. (1974) Determinants of bid-ask spreads in the over-thecounter market. *Journal of Financial Economics,* 1**,** 353-364.

BENSTON ROBERT, L. & GEORGE, J. (1974) Determinants of bid-asked spreads in the over-the-counter market* 1. *Journal of Financial Economics,* 1**,** 353-364.

BENZ, E. & TRÜCK, S. (2006) CO2 Emission Allowances Trading in Europe-Specifying a new Class of Assets. *Problems and Perspectives in Management,* 4**,** 30-40.

BENZ, E. & TRÜCK, S. (2009) Modeling the price dynamics of CO2 emission allowances. *Energy Economics,* 31**,** 4-15.

BENZ, E. A. & HENGELBROCK, J. (2008) Price Discovery and Liquidity in the European CO2 Futures Market: An Intraday Analysis. *Electronic copy available at: http://ssrn.com/abstract=1131770*.

BERKMAN, H., BRAILSFORD, T. & FRINO, A. (2005) A note on execution costs for stock index futures: Information versus liquidity effects. *Journal of Banking & Finance,* 29**,** 565-577.

BERNARD, B. S. Autocorrelation of the Trade Process: Evidence from the Helsinki Stock Exchange. *The Quarterly Review of Economics and Finance*.

BERRY, T. D. & HOWE, K. M. (1994) Public information arrival. American Finance Association; Blackwell Publishing.

BERTRAM, W. K. (2005) A threshold model for Australian Stock Exchange equities. Elsevier.

BHATTACHARYA, S. & MOOKHERJEE, D. (1986) Portfolio choice in research and development. *The RAND Journal of Economics*, 17**,** 594-605.

BIAIS, B. (1993) Price formation and equilibrium liquidity in fragmented and centralized markets. *Journal of Finance,* 48**,** 157-185.

BIAIS, B., GLOSTEN, L. & SPATT, C. (2005) Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets,* 8**,** 217-264.

BIAIS, B., HILLION, P. & SPATT, C. (1995) An empirical analysis of the limit order book and the order flow in the Paris Bourse. *The Journal of Finance,* 50**,** 1655-1689.

BIAIS, B., HILLION, P. & SPATT, C. (1999) Price discovery and learning during the preopening period in the Paris Bourse. *The Journal of Political Economy,* 107**,** 1218-1248.

BIKHCHANDANI, S., HIRSHLEIFER, D. & WELCH, I. (1992) A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades, 100J. *Pol. Econ,* 992**,** 994.

BLACK, F. (1986) Noise. *Journal of finance,* 41**,** 529-43.

BLOOMFIELD, R. & O'HARA, M. (1999) Market transparency: Who wins and who loses? *Review of Financial Studies,* 12**,** 5.

BLUME, L., EASLEY, D. & O'HARA, M. (1994) Market statistics and technical analysis: The role of volume. *Journal of Finance,* 49**,** 153-181.

BLUME, L. E. & EASLEY, D. (1984) Rational expectations equilibrium: An alternative approach* 1. *Journal of Economic Theory,* 34**,** 116-129.

BLUME, L. E. & EASLEY, D. (1998) Rational expectations and rational learning. *Organizations with incomplete information: Essays in economic analysis: A tribute to Roy Radner***,** 61–109.

BOARD, J. & SUTCLIFFE, C. (1995) The effects of trade transparency in the London Stock Exchange: A summary. *LSE Financial Markets Group Special Paper*.

BOEHMER, E., GRAMMIG, J. & THEISSEN, E. (2007) Estimating the probability of informed trading--does trade misclassification matter? *Journal of Financial Markets,* 10**,** 26-47.

BÖHRINGER, C. & LANGE, A. (2005a) Economic Implications of Alternative Allocation Schemes for Emission Allowances. *Scandinavian Journal of Economics,* 107**,** 563-581.

BÖHRINGER, C. & LANGE, A. (2005b) Mission Impossible!? On the Harmonization of National Allocation Plans under the EU Emissions Trading Directive. *Journal of Regulatory Economics,* 27**,** 81-94.

BÖHRINGER, C. & LANGE, A. (2005) On the design of optimal grandfathering schemes for emission allowances. *European Economic Review,* 49**,** 2041-2055.

BOLLEN, N. P. B., SMITH, T. & WHALEY, R. E. (2004) Modeling the bid/ask spread: measuring the inventory-holding premium* 1. *Journal of Financial Economics,* 72**,** 97-141.

BOLLEN, N. P. B. & WHALEY, R. E. (2004) Does net buying pressure affect the shape of implied volatility functions? *The Journal of Finance,* 59**,** 711-753.

BOLLERSLEV, T. (1986) Generalized autoregressive conditional heteroskedasticity.

BOLLERSLEV, T. (1990) Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The Review of Economics and Statistics,* 72**,** 498-505.

BOLLERSLEV, T., CHOU, R. Y. & KRONER, K. F. (1992) ARCH modeling in finance* 1:: A review of the theory and empirical evidence. *Journal of econometrics,* 52**,** 5-59.

BOLLERSLEV, T. & DOMOWITZ, I. (1991) Price volatility, spread variability, and the role of alternative market mechanisms. *Review of Futures Markets,* 10**,** 78–102.

BOLLERSLEV, T. & DOMOWITZ, I. (1993) Trading patterns and prices in the interbank foreign exchange market. *The Journal of Finance,* 48**,** 1421-1443.

BOLLERSLEV, T., ENGLE, R. F. & NELSON, D. B. (1995) ARCH models, in 'Handbook of Econometrics', Vol. 4. North-Holland, New York, NY.

BOLLERSLEV, T., LITVINOVA, J. & TAUCHEN, G. (2006) Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics,* 4**,** 353.

BOLLERSLEV, T. & MELVIN, M. (1994) Bid--ask spreads and volatility in the foreign exchange market:: An empirical analysis. *Journal of International Economics,* 36**,** 355-372.

BOLLERSLEV, T. & WOOLDRIDGE, J. M. (1992) Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews,* 11**,** 143-172.

BORAK, S., HÄRDLE, W., TRÜCK, S. & WERON, R. (2006) Convenience yields for CO2 emission allowance futures contracts. *SFB 649 Discussion Papers*.

BOUCHAUD, J. P., GEFEN, Y., POTTERS, M. & WYART, M. (2004) Fluctuations and response in financial markets: the subtle nature of 'random price changes. *Quantitative Finance,* 4**,** 176-190.

BOUCHAUD, J. P., KOCKELKOREN, J. & POTTERS, M. (2006) Random walks, liquidity molasses and critical response in financial markets. *Quantitative Finance,* 6**,** 115-123.

BOUGEROL, P. & PICARD, N. (1992) Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics,* 52**,** 115-127.

BOWE, M., HYDE, STUART AND MCFARLANE, LAVERN, (2007) Duration, Trading Volume and the Price Impact of Trades in an Emerging Futures Market *Manchester Business School Working Paper No. 519. .*

BOX, G. E. P. & COX, D. R. (1964) An analysis of transformations. Blackwell Publishing; Royal Statistical Society.

BOX-STEFFENSMEIER, J. M. & JONES, B. S. (2004) *Event history modeling: a guide for social scientists,* Cambridge University Press.

BRADFIELD, J. (1979) A formal dynamic model of market making. *Journal of Financial and Quantitative Analysis,* 14**,** 275-291.

BRENNAN, M. J. & SUBRAHMANYAM, A. (1996) Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics,* 41**,** 441-464.

BROCK, W., LAKONISHOK, J. & LEBARON, B. (1992) Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance,* 47**,** 1731-1764.

BROCK, W. & LEBARON, B. D. (1995) A dynamic structural model for stock return volatility and trading volume. National Bureau of Economic Research Cambridge, Mass., USA.

BROCK, W. A. (2001) Asset price behavior in complex environments. *Growth theory, nonlinear dynamics, and economic modelling: scientific essays of William Allen Brock***,** 363.

BROCK, W. A. & HOMMES, C. H. (1997) A rational route to randomness. *Econometrica,* 65**,** 1059-1095.

BROCK, W. A. & KLEIDON, A. W. (1992) Periodic market closure and trading volume:: A model of intraday bids and asks. *Journal of Economic Dynamics and Control,* 16**,** 451-489.

BROCKMAN, P. & CHUNG, D. Y. (1998) Inter-and intra-day liquidity patterns on the Stock Exchange of Hong Kong. *Journal of International Financial Markets, Institutions and Money,* 8**,** 277-298.

BROOKS, C. (1997) Linear and non-linear (non-) forecastability of high-frequency exchange rates.

BROWN, D. P. & ZHANG, Z. M. (1997) Market orders and market efficiency. *Journal of Finance,* 52**,** 277-308.

BRUNNERMEIER, M. K. (2001) *Asset pricing under asymmetric information: bubbles, crashes, technical analysis, and herding,* Oxford University Press.

BUNN, D. & FEZZI, C. (2007) Interaction of European carbon trading and energy prices. *Fondazione Eni Enrico Mattei Working Papers***,** 123.

BURDETT, K. & O'HARA, M. (1987) Building blocks* 1:: An introduction to block trading. *Journal of Banking & Finance,* 11**,** 193-212.

BURTRAW, D., PALMER, K., BHARVIRKAR, R. & PAUL, A. (2002) The effect on asset values of the allocation of carbon dioxide emission allowances. *The Electricity Journal,* 15**,** 51-62.

CALAMIA, A. (1999) Market Microstructure: Theory and Empirics. *LEM Papers Series,* 19.

CAMINAL, R. & VIVES, X. (1996) Why market shares matter: an information-based theory. *The RAND Journal of Economics***,** 221-239.

CAMINAL, R. & VIVES, X. (1997) DP1744 Price Dynamics and Consumer Learning.

CAMPBELL, J. Y., LO, A. W., MACKINLAY, A. C. & WHITELAW, R. F. (1998) The econometrics of financial markets. Cambridge Univ Press.

CAPOOR, K. & AMBROSI, P. (2006) State and Trends of the Carbon Market 2006. World Bank Washington, DC.

CAPOOR, K. & AMBROSI, P. (2007) State and Trends of the Carbon Market 2007.

CAPOOR, K. & AMBROSI, P. (2008) State and Trends of the Carbon Market 2008.

CAPOOR, K. & AMBROSI, P. (2009) State and Trends of the Carbon Market 2009.

CARMONA, M., LOPES, R. I., BORBA, M., OMOKAWA, M., NAUFAL, R., MIYAJI, K., MATSUMURA, N., VIEIRA, N. & PEREIRA, P. R. B. (2008) Comparison of the effects of carbon dioxide and helium pneumoperitoneum on renal function. *Journal of Endourology,* 22**,** 1077-1082.

CARMONA, R. (2009) *Indifference pricing: theory and applications,* Princeton Univ Pr.

CARMONA, R. & LUDKOVSKI, M. (2008) Pricing asset scheduling flexibility using optimal switching. *Applied Mathematical Finance,* 15**,** 405-447.

CHAKRAVARTY, S. & HOLDEN, C. W. (1995) An integrated model of market and limit orders. *Journal of Financial Intermediation,* 4**,** 213-241.

CHAMLEY, C. (2003) Dynamic speculative attacks. *The American Economic Review,* 93**,** 603-621.

CHAMLEY, C. (2003) *Learning and Speculation Against a Fixed Exchange Rate.* Cambridge University Press.

CHAMLEY, C. (2003) *Multiple levels of participation and learning in financial markets.* Cambridge University Press.

CHAMLEY, C. (2004) *Rational herds: economic models of social learning,* Cambridge University Press.

CHAN, K. (1992) A further analysis of the lead-lag relationship between the cash market and stock index futures market. *Review of Financial Studies,* 5**,** 123.

CHAN, K. & FONG, W. M. (2000) Trade size, order imbalance, and the volatility-volume relation* 1. *Journal of Financial Economics,* 57**,** 247-273.

CHAN, L. K. C. & LAKONISHOK, J. (1993) Institutional trades and intraday stock price behavior. *Journal of Financial Economics,* 33**,** 173-199.

CHAN, Y. E. (2000) IT value: The great divide between qualitative and quantitative and individual and organizational measures. *Journal of Management Information Systems,* 16**,** 225-261.

CHEVALLIER, J. (2009) Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Economics,* 31**,** 614-625.

CHEVALLIER, J. (2009) Energy risk management with carbon assets. *International Journal of Global Energy Issues,* 32**,** 328-349.

CHEVALLIER, J. (2009) A Note on Cointegrating and Vector Autoregressive Relationships between CO2 allowances spot and futures prices. *Economics Bulletin,* 30**,** 1564-1584.

CHEVALLIER, J. (2010) EUAs and CERs: Vector autoregression, impulse response function and cointegration analysis. *Economics Bulletin,* 30**,** 558-576.

CHEVALLIER, J. (2010) Volatility forecasting of carbon prices using factor models". *Economics Bulletin,* 30**,** 1642-1660.

CHEVALLIER, J. & SÉVI, B. Jump-robust estimation of realized volatility in the EU emissions trading scheme. *Journal of Energy Markets,* 3**,** 49.

CHIARELLA, C. & NOVIKOV, A. *Contemporary Quantitative Finance: Essays in Honour of Eckhard Platen*, Springer.

CHIAROMONTE, F. & DOSI, G. (1998) Modeling a Decentralized Asset Market: An Introduction to the Financial "Toy-Room". *Interim Report, IIASA, Laxemburg, Austria*.

CHOI, J. Y., SALANDRO, D. & SHASTRI, K. (1988) On the estimation of bid-ask spreads: Theory and evidence. *Journal of Financial and Quantitative Analysis,* 23**,** 219-230.

CHRIS LEACH, J. & MADHAVAN, A. N. (1992) Intertemporal price discovery by market makers: Active versus passive learning. *Journal of Financial Intermediation,* 2**,** 207-235.

CHRISTIANSEN, A. C., ARVANITAKIS, A., TANGEN, K. & HASSELKNIPPE, H. (2005) Price determinants in the EU emissions trading scheme. *Climate Policy,* 5**,** 15-30.

CHRISTIE, W. G. & SCHULTZ, P. H. (1994) Why do NASDAQ market makers avoid odd-eighth quotes? *Journal of Finance,* 49**,** 1813-1840.

CHRISTOFFERSEN, P. & PELLETIER, D. (2004) Backtesting value-at-risk: A duration-based approach. Oxford Univ Press.

CLEMENTS, M. P. & HENDRY, D. F. (1998) *Forecasting economic time series*, Cambridge Univ Pr.

COHEN, K. J., HAWAWINI, G. A., MAIER, S. F., SCHWARTZ, R. A. & WHITCOMB, D. K. (1980) Implications of microstructure theory for empirical research on stock price behavior. *Journal of Finance***,** 249-257.

COHEN, K. J., MAIER, S. F., SCHWARTZ, R. A. & WHITCOMB, D. K. (1981) Transaction costs, order placement strategy, and existence of the bid-ask spread. *The Journal of Political Economy,* 89**,** 287-305.

CONRAD, C. (2008) Non-negativity conditions for the hyperbolic GARCH model. *Journal of Econometrics,* 157**,** 441-457.

CONRAD, C. & KARANASOS, M. Negative volatility spillovers in the unrestricted ECCC-GARCH model. *Econometric Theory,* 26**,** 838-862.

CONRAD, C., KARANASOS, M. & ZENG, N. (2010) Multivariate fractionally integrated APARCH modeling of stock market volatility: A multi-country study. *Journal of Empirical Finance*.

CONRAD, C. & LAMLA, M. J. (2010) The High Frequency Response of the EUR USD Exchange Rate to ECB Communication. *Journal of Money, Credit and Banking,* 42**,** 1391-1417.

CONRAD, J. M. *Resource economics*, Cambridge Univ Pr.

CONVERY, F. J., ELLERMAN, A. D. & PERTHUIS DE, C. (2008) The European Carbon Market in Action: Lessons from the first Trading Period. *Interim Report, IIASA, Laxemburg, Austria*.

CONVERY, F. J. & REDMOND, L. (2007) Market and price developments in the European Union emissions trading scheme. *Review of Environmental Economics and Policy,* 1**,** 88.

COPELAND, T. E. & GALAI, D. (1983) Information effects on the bid-ask spread. *Journal of finance,* 38**,** 1457-1469.

CORSI, F., MITTNIK, S., PIGORSCH, C. & PIGORSCH, U. (2008) The volatility of realized volatility. *Econometric Reviews,* 27**,** 46-78.

CORSI, T. M., BOYSON, S., VERBRAECK, A., VAN HOUTEN, S., HAN, C. & MACDONALD, J. R. (2006) The real-time global supply chain game: New educational tool for developing supply chain management professionals. *Transportation Journal,* 45**,** 61-73.

CURCIO, R. & GOODHART, C. (1992) When support/resistance levels are broken, can profits be made? Evidence from the foreign exchange market. *FMG Discussion Papers*.

CUTLER, D. M., POTERBA, J. M. & SUMMERS, L. H. (1990) Speculative dynamics. National Bureau of Economic Research Cambridge, Mass., USA.

CZADO, C. & HAUG, S. An ACD-ECOGARCH (1, 1) Model. *Journal of Financial Econometrics,* 8**,** 335.

DACOROGNA, M. M. (2001) *An introduction to high-frequency finance*, Academic Press.

DACOROGNA, M. M., MÜLLER, U. A., NAGLER, R. J., OLSEN, R. B. & PICTET, O. V. (1993) A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance,* 12**,** 413-438.

DASKALAKIS, C., MEHTA, A. & PAPADIMITRIOU, C. (2006) A note on approximate Nash equilibria. *Internet and Network Economics***,** 297-306.

DASKALAKIS, G. & MARKELLOS, R. N. (2008) Are the European Carbon Markets efficient? *Review of Futures Markets,* 17**,** 103-128.

DASKALAKIS, G., PSYCHOYIOS, D. & MARKELLOS, R. N. (2009) Modeling CO2 emission allowance prices and derivatives: Evidence from the European trading scheme. *Journal of Banking & Finance,* 33**,** 1230-1241.

DE JONG, F. (2002) Measures of contributions to price discovery: a comparison* 1. *Journal of Financial Markets,* 5**,** 323-327.

DE JONG, F., NIJMAN, T. & ROELL, A. (1996) Price effects of trading and components of the bid-ask spread on the Paris Bourse. *Journal of Empirical Finance,* 3**,** 193-213.

DE JONG, F., NIJMAN, T. & RÖELL, A. (1995) A comparison of the cost of trading French shares on the Paris Bourse and on SEAQ International. *European Economic Review,* 39**,** 1277-1301.

DE JONG, F., NIJMAN, T. & RÖELL, A. (1996) Price effects of trading and components of the bid-ask spread on the Paris Bourse. *Journal of Empirical Finance,* 3**,** 193-213.

DE JONG, G. C. (1990) An indirect utility model of car ownership and private car use. *European Economic Review,* 34**,** 971-985.

DE LONG, J. B., SHLEIFER, A., SUMMERS, L. H. & WALDMANN, R. J. (1990) Noise trader risk in financial markets. *Journal of Political Economy,* 98**,** 703-738.

DE LUCA, G. & GALLO, G. M. (2004) Mixture processes for financial intradaily durations. *Studies in Nonlinear Dynamics & Econometrics,* 8**,** 1223.

DE LUCA, G. & GALLO, G. M. (2009) Time-varying mixing weights in mixture autoregressive conditional duration models. Taylor & Francis Group, 4 Park Square, Milton Park, Abingdon, Oxon, OX 14 4 RN, UK.

DE LUCA, G., RIVIECCIO, G. & ZUCCOLOTTO, P. (2008) Exploring the Copula Approach for the Analysis of Financial Durations. Springer.

DE LUCA, G., RIVIECCIO, G. & ZUCCOLOTTO, P. (2009) ARCHIMEDEAN COPULAE AND MARKET RISK IN A FINANCIAL CRISIS PERSPECTIVE.

DE LUCA, G. & ZUCCOLOTTO, P. (2003) Finite and infinite mixtures for financial durations. Dipartimento di Statistica, Probabilit e Statistiche Applicate-University of Rome.

DE LUCA, G. & ZUCCOLOTTO, P. (2006) Regime-switching Pareto distributions for ACD models. Elsevier.

DEB. HARRIS, F. H., MCINISH, T. H., SHOESMITH, G. L. & WOOD, R. A. (1995) Cointegration, error correction, and price discovery on informationally linked security markets. *Journal of financial and quantitative analysis,* 30**,** 563-579.

DECLERCK, F. (2000) *Analyse des meilleures limites du carnet d'ordres: application à la Bourse de Paris*, ANRT, Université Pierre Mendes France (Grenoble II).

DEGRYSE, H. (1999) The total cost of trading Belgian shares: Brussels versus London. *Journal of Banking & Finance,* 23**,** 1331-1355.

DEGRYSE, H. & VANACHTER, M. (2001) *Alternative trading systems and liquidity*, Citeseer.

DEGRYSE, H. A. (1999) Essays on financial intermediation, product differentiation and market structure. *Revue de la Banque/Bank-en Financiewezen***,** 78-81.

DEMOS, A. A. & GOODHART, C. A. E. (1996) The interaction between the frequency of market quotations, spread and volatility in the foreign exchange market. *Applied Economics,* 28**,** 377-386.

DEMSETZ, H. (1968) THE COST OF TRANSACTING. *The Quarterly Journal of Economics,* 82**,** 33-53.

DEMSETZ, H. (1968) The cost of transacting. The MIT Press.

DEMSETZ, H. (1971) On the regulation of industry: a reply. *The Journal of Political Economy,* 79**,** 356-363.

DEREK, W. B. & FEZZI, C. (2007) Interaction of European Carbon Trading and Energy Prices. *Fondazione Eni Enrico Mattei Working paper*, 123.

DIAMOND, D. W. & VERRECCHIA, R. E. (1987) Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics,* 18**,** 277-311.

DICKEY, D. A. & FULLER, W. A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association,* 74**,** 427-431.

DIEBOLD, F. X. & LOPEZ, J. A. (1996) Forecast evaluation and combination.

DIEBOLD, F. X. & MARIANO, R. S. (1995) Comparing predictive accuracy. American Statistical Association.

DOLADO, J. J., RODRIGUEZ-POO, J. & VEREDAS, D. (2004) Testing weak exogeneity in the exponential family: an application to financial point processes. Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

DOMOWITZ, I. (1990) The mechanics of automated trade execution systems* 1. *Journal of Financial Intermediation,* 1**,** 167-194.

DROST, F. C. & WERKER, B. J. M. (1996) Closing the GARCH gap: Continuous time GARCH modeling. *Journal of Econometrics,* 74**,** 31-57.

DUFOUR, A. & ENGLE, R. F. (2000a) The ACD Model: Predictability of the TIme Between Consecutive Trades. *Discussion Papers in Finance: ISMA Centre*.

DUFOUR, A. & ENGLE, R. F. (2000b) Time and the price impact of a trade. *The Journal of Finance,* 55**,** 2467-2498.

DURBIN, M. *All About High-Frequency Trading*, McGraw-Hill.

DUTTA, P. K. & MADHAVAN, A. (1997) Competition and collusion in dealer markets. *Journal of Finance***,** 245-276.

EASLEY, D., ENGLE, R. F., O'HARA, M. & WU, L. (2008) Time-varying arrival rates of informed and uninformed trades. *Journal of Financial Econometrics,* 6**,** 171.

EASLEY, D., HVIDJKAER, S. & O'HARA, M. (2002) Is information risk a determinant of asset prices? *Journal of Finance,* 57**,** 2185-2221.

EASLEY, D., KIEFER, N. M. & O'HARA, M. (1997) The information content of the trading process* 1. *Journal of Empirical Finance,* 4**,** 159-186.

EASLEY, D., KIEFER, N. M. & O'HARA, M. (1997) One day in the life of a very common stock. *Review of Financial Studies,* 10**,** 805.

EASLEY, D., KIEFER, N. M., O'HARA, M. & PAPERMAN, J. B. (1996) Liquidity, information, and infrequently traded stocks. *Journal of Finance,* 51**,** 1405-1436.

EASLEY, D. & O'HARA, M. (1987) Price, trade size, and information in securities markets* 1. *Journal of Financial economics,* 19**,** 69-90.

EASLEY, D. & O'HARA, M. (1992) Time and the process of security price adjustment. *Journal of Finance,* 47**,** 577-605.

EASLEY, D. & O'HARA, M. (2003) Microstructure and asset pricing. *Handbook of the Economics of Finance,* 1**,** 1021-1051.

EASLEY, D. & O'HARA, M. (2009) Ambiguity and nonparticipation: the role of regulation. *Review of Financial Studies*.

EASLEY, D., O'HARA, M. & SAAR, G. (2001) How stock splits affect trading: A microstructure approach. *Journal of Financial and Quantitative Analysis,* 36**,** 25-51.

EASLEY, D., O'HARA, M. & SRINIVAS, P. S. (1998) Option volume and stock prices: Evidence on where informed traders trade. *The Journal of Finance,* 53**,** 431-465.

ELLERMAN, A. D. (2005) US Experience with Emissions Trading. *Emissions Trading for Climate Policy***,** 78–95.

ELLERMAN, A. D. & BUCHNER, B. K. (2007) The European union emissions trading scheme: Origins, allocation, and early results.

ELLERMAN, A. D. & BUCHNER, B. K. (2008) Over-allocation or abatement? A preliminary analysis of the EU ETS based on the 2005–06 emissions data. *Environmental and Resource Economics,* 41**,** 267-287.

ENGLE, C. R. & STONE, N. M. (1997) *Developing Business Proposals for Aquaculture Loans*, Southern Regional Aquaculture Center.

ENGLE III, R. F., ITO, T. & LIN, W. L. (1991) Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market. National Bureau of Economic Research Cambridge, Mass., USA.

ENGLE, R. (2002) New frontiers for ARCH models. *Journal of Applied Econometrics,* 17**,** 425-446.

ENGLE, R. F. (2000) The econometrics of ultra-high-frequency data. *Econometrica,* 68**,** 1-22.

ENGLE, R. F. & BOLLERSLEV, T. (1986) Modelling the persistence of conditional variances. *Econometric reviews,* 5**,** 1-50.

ENGLE, R. F. & GONZALEZ-RIVERA, G. (1991) Semiparametric ARCH models. *Journal of Business & Economic Statistics,* 9**,** 345-359.

ENGLE, R. F. & GRANGER, C. W. J. (1987) Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society***,** 251-276.

ENGLE, R. F., HENDRY, D. F. & RICHARD, J. F. (1983) Exogeneity. *Econometrica: Journal of the Econometric Society***,** 277-304.

ENGLE, R. F. & LANGE, J. (2001) Predicting VNET: A model of the dynamics of market depth 1. *Journal of Financial Markets,* 4**,** 113-142.

ENGLE, R. F. & LUNDE, A. (2003) Trades and quotes: a bivariate point process. *Journal of Financial Econometrics,* 1**,** 159.

ENGLE, R. F. & PATTON, A. J. (2004) Impacts of trades in an error-correction model of quote prices 1. *Journal of Financial Markets,* 7**,** 1-25.

ENGLE, R. F. & RUSSELL, J. (1994) Forecasting transaction rates: the autoregressive conditional duration model.

ENGLE, R. F. & RUSSELL, J. R. (1997) Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model* 1. *Journal of Empirical Finance,* 4**,** 187-212.

ENGLE, R. F. & RUSSELL, J. R. (1998) Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica,* 66**,** 1127-1162.

ENGLE, R. F. & RUSSELL, J. R. (2004) Analysis of high frequency data. *Handbook of Financial Econometrics*.

EPPS, T. W. (2009) *Quantitative finance: its development, mathematical foundations, and current scope*, Wiley.

FAMA, E. F. (1965) The behavior of stock-market prices. *Journal of business,* 38**,** 34-105.

FAMA, E. F. (1970) Efficient capital markets: A review of theory and empirical work. *The journal of Finance,* 25**,** 383-417.

FELDMAN, P. (1987) A practical scheme for non-interactive verifiable secret sharing. IEEE.

FELDMAN, R. D. (1978) Rationing congested goods by preferences on quality. *Public Finance= Finances publiques,* 33**,** 225.

FERNANDES, M. & GRAMMIG, J. (2006) A family of autoregressive conditional duration models. *Journal of Econometrics,* 130**,** 1-23.

FLETCHER, J. (1995) An examination of the selectivity and market timing performance of UK unit trusts. *Journal of Business Finance & Accounting,* 22**,** 143-156.

FLOOD, M. D., HUISMAN, R., KOEDIJK, K. G. & MAHIEU, R. J. (1999) Quote disclosure and price discovery in multiple-dealer financial markets. *Review of Financial Studies,* 12**,** 37.

FLOOD, P. (1998) Is HRM dead? What will happen to HRM when traditional methods are gone? *Human Resource Management: The New Agenda, Pitman, London***,** 52-71.

FOSTER, F. D. & VISWANATHAN, S. (1990) A theory of the interday variations in volume, variance, and trading costs in securities markets. *Review of Financial Studies,* 3**,** 593.

FOSTER, F. D. & VISWANATHAN, S. (1993) The effect of public information and competition on trading volume and price volatility. *Review of Financial Studies,* 6**,** 23.

FOSTER, F. D. & VISWANATHAN, S. (1993) Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models. *Journal of Finance,* 48**,** 187-211.

FOSTER, F. D. & VISWANATHAN, S. (1996) Strategic trading when agents forecast the forecasts of others. *Journal of Finance,* 51**,** 1437-1478.

FOUCAULT, T. (1999) Order flow composition and trading costs in a dynamic limit order market1. *Journal of Financial markets,* 2**,** 99-134.

FOUCAULT, T., KADAN, O. & KANDEL, E. (2005) Limit order book as a market for liquidity. *Review of Financial Studies,* 18**,** 1171.

FOUCAULT, T., ROELL, A. & SANDAS, P. (2003) Market making with costly monitoring: An analysis of the SOES controversy. *Review of Financial Studies,* 16**,** 345.

FRENCH, K. & ROLL, R. (1986) Stock Market Variances: the Arrival of Information and the Reaction of Traders,". *Journal of Financial Economics,* 17**,** 5–26.

FRENCH, K. & ROLL, R. (1986) Stock return variances: The arrival of information and the reaction of traders. *Journal of Financial Economics,* 17**,** 5-26.

FRENCH, K. R. (1980) Stock returns and the weekend effect* 1. *Journal of Financial Economics,* 8**,** 55-69.

FRINO, A., KRUK, J. & LEPONE, A. (2010) Liquidity and transaction costs in the European carbon futures market. *Journal of Derivatives & Hedge Funds,* 16**,** 100-115.

FRYDMAN, R. (1982) Towards an understanding of market processes: individual expectations, learning, and convergence to rational expectations equilibrium. *The American Economic Review***,** 652-668.

FUDENBERG, D. & KREPS, D. (1988) Learning, experimentation and equilibrium in games. *photocopy, Department of Economics, Stanford University*.

GABRIÉ, H. & JACQUIER, J. L. (1994) *La théorie moderne de l'entreprise: l'approche institutionnelle*, Economica.

GALLANT, A. R., ROSSI, P. E. & TAUCHEN, G. (1992) Stock prices and volume. *Review of Financial studies,* 5**,** 199.

GARBADE, K. D. & SILBER, W. L. (1979) Structural organization of secondary markets: Clearing frequency, dealer activity and liquidity risk. *Journal of Finance***,** 577-593.

GARMAN, M. B. (1976) Market microstructure* 1. *Journal of Financial Economics,* 3**,** 257-275.

GEMMILL, G. (1996) Transparency and liquidity: A study of block trades on the London Stock Exchange under different publication rules. *The Journal of Finance,* 51**,** 1765-1790.

GENNOTTE, G. & LELAND, H. (1990) Market liquidity, hedging, and crashes. *The American Economic Review,* 80**,** 999-1021.

GEORGE, T. J., KAUL, G. & NIMALENDRAN, M. (1991) Estimation of the bid-ask spread and its components: A new approach. *Review of Financial Studies,* 4**,** 623.

GERETY, M. S. & MULHERIN, J. H. (1991) Patterns in intraday stock market volatility, past and present. *Financial Analysts Journal***,** 71-79.

GERETY, M. S. & MULHERIN, J. H. (1992) Trading halts and market activity: an analysis of volume at the open and the close. *The Journal of Finance,* 47**,** 1765-1784.

GERETY, M. S. & MULHERIN, J. H. (1994) Price formation on stock exchanges: The evolution of trading within the day. *Review of Financial Studies,* 7**,** 609.

GERHARD, F. & HAUTSCH, N. (2000) *Determinants of inter-trade durations and hazard rates using proportional hazard ARMA models*, Universität Konstanz, Fachbereich für Wirtschaftswissenschaften.

GERHARD, F. & HAUTSCH, N. (2007) A dynamic semiparametric proportional hazard model. *Studies in Nonlinear Dynamics & Econometrics,* 11**,** 1.

GHYSELS, E., GOURIEROUX, C. & JASIAK, J. (2004) Stochastic volatility duration models. *Journal of Econometrics,* 119**,** 413-433.

GHYSELS, E., GOURIÉROUX, C., JASIAK, J. & CIRANO (1995) *Trading patterns, time deformation and stochastic volatility in foreign exchange markets*, CIRANO.

GHYSELS, E., HARVEY, A. C. & RENAULT, E. (1996) Stochastic volatility, chapter 14 in GS Maddala and CR Rao (eds), Handbook of Statistics, Vol. 14. North-Holland, Amsterdam.

GHYSELS, E. & JASIAK, J. (1998) GARCH for irregularly spaced financial data: the ACD-GARCH model. *Studies in Nonlinear Dynamics and Econometrics,* 2**,** 133-149.

GIBBONS, M. R. & HESS, P. (1981) Day of the week effects and asset returns. *Journal of Business,* 54**,** 579-596.

GIOT, P. (2000) Time transformations, intraday data, and volatility models. *Journal of Computational Finance,* 4**,** 31-62.

GIOT, P. (2005) Market risk models for intraday data. *The European Journal of Finance,* 11**,** 309-324.

GIOT, P. & UNIVERSITÉ CATHOLIQUE DE LOUVAIN . CENTER FOR OPERATIONS RESEARCH AND, E. (1999) *Time transformations, intraday data and volatility models,* Citeseer.

GLOSTEN, L. & MILGROM, P. (1985) Bid, Ask and Transactions Prices in a Specialist Market with Insider Trading. *Journal of Financial Economics,* 14**,** 71-100.

GLOSTEN, L. R. (1987) Components of the bid-ask spread and the statistical properties of transaction prices. *Journal of Finance,* 42**,** 1293-1307.

GLOSTEN, L. R. (1989) Insider trading, liquidity, and the role of the monopolist specialist. *Journal of Business*, 62**,** 211-235.

GLOSTEN, L. R. (1994) Is the electronic open limit order book inevitable? *Journal of Finance,* 49**,** 1127-1161.

GLOSTEN, L. R. (1998) Competition, design of exchanges and welfare. working paper, Columbia University.

GLOSTEN, L. R. & HARRIS, L. E. (1988) Estimating the components of the bid/ask spread. *Journal of financial Economics,* 21**,** 123-142.

GLOSTEN, L. R. & MILGROM, P. (1985) Bid, ask and transaction prices in a market-maker market with heterogeneously informed traders. *Journal of Financial Economics,* 14**,** 71–100.

GLOSTEN PAUL, R. & LAWRENCE, R. (1985) Bid, ask and transaction prices in a specialist market with heterogeneously informed traders* 1. *Journal of financial economics,* 14**,** 71-100.

GLOSTEN, R. (1994) A contingent claim approach to performance evaluation. *Journal of Empirical Finance,* 1**,** 133-160.

GODAL, O. & KLAASSEN, G. (2006) Carbon trading across sources and periods constrained by the Marrakesh Accords. *Journal of Environmental Economics and Management*, 51**,** 308-322.

GOETTLER, R. L. & LESLIE, P. (2005) Cofinancing to manage risk in the motion picture industry. *Journal of Economics & Management Strategy,* 14**,** 231-261.

GOETTLER, R. L., PARLOUR, C. A. & RAJAN, U. (2005) Equilibrium in a dynamic limit order market. *The Journal of Finance,* 60**,** 2149-2192.

GOLDMAN, M. B. & BEJA, A. (1979) Market prices vs. equilibrium prices: Returns' variance, serial correlation, and the role of the specialist. *The Journal of Finance,* 34**,** 595-607.

GONZALO, J. & GRANGER, C. (1995) Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics,* 13**,** 27-35.

GOODHART, C. (1990) News and the foreign exchange market. *FMG Discussion Papers*.

GOODHART, C. & DEMOS, A. (1991) Reuters’ screen images of the foreign exchange markets: The Yen/Dollar and the Sterling/Dollar spot market. *Journal of International Securities Markets,* 5**,** 35–64.

GOODHART, C., ITO, T. & PAYNE, R. (1996) One day in June 1993: A study of the working of the Reuters 2000-2 electronic foreign exchange trading system. University of Chicago Press.

GOODHART, C. A. E. & FIGLIUOLI, L. (1991) Every minute counts in financial markets. *Journal of International Money and Finance,* 10**,** 23-52.

GOODHART, C. A. E. & FIGLIUOLI, L. (1992) The geographical location of the foreign exchange market: a test of an “islands” hypothesis. *Journal of International and Comparative Economics,* 1**,** 13-28.

GOODHART, C. A. E. & GIUGALE, M. (1993) From hour to hour in the foreign exchange market. *The Manchester School,* 61**,** 1-34.

GOODHART, C. A. E., HALL, S. G., HENRY, S. G. B. & PESARAN, B. (1993) News effects in a high frequency model of the sterling dollar exchange rate. *Journal of Applied Econometrics,* 8**,** 1-13.

GOODHART, C. A. E. & O'HARA, M. (1997) High frequency data in financial markets: Issues and applications* 1. *Journal of Empirical Finance,* 4**,** 73-114.

GOODHART RICHARD, G. & CHARLES, A. E. (1996) Microstructural dynamics in a foreign exchange electronic broking system. *Journal of International Money and Finance,* 15**,** 829-852.

GOURIÉROUX, C., JASIAK, J. & LEFOL, G. (1999) Intraday trading activity. *Journal of Financial Markets,* 2**,** 193-226.

GOURIEROUX, C., JASIAK, J. & SUFANA, R. (2004) A Dynamic Model for Multivariate Stochastic Volatility: The Wishart Autoregressive Process. Working Paper.

GRAMMIG, J. & MAURER, K. O. (2000) Non monotonic hazard functions and the autoregressive conditional duration model. *Econometrics Journal,* 3**,** 16-38.

GRAMMIG, J. & WELLNER, M. (2002) Modeling the interdependence of volatility and inter-transaction duration processes. *Journal of Econometrics,* 106**,** 369-400.

GRAMMIG, J., WUENSCHE, O. & THEISSEN, E. (2007) Time and the Price Impact of a Trade: A Structural Approach.

GRANGER, C. W. J. (1981) Some properties of time series data and their use in econometric model specification. *Journal of econometrics,* 16**,** 121-130.

GREENE, W. H. (2002) LIMDEP version 8.0: Econometric modeling guide.

GROSSMAN, S. J. (1992) The informational role of upstairs and downstairs trading. *The Journal of Business,* 65**,** 509-528.

GROSSMAN, S. J. & STIGLITZ, J. E. (1980) On the impossibility of informationally efficient markets. *The American Economic Review,* 70**,** 393-408.

GROSSMAN, S. J. & STIGLITZ, J. E. (1980) Stockholder unanimity in making production and financial decisions. *The Quarterly Journal of Economics,* 94**,** 543-566.

GRUBB, M., LAING, T., COUNSELL, T. & WILLAN, C. Global carbon mechanisms: lessons and implications. *Climatic Change***,** 1-35.

GUILLAUME, D. M., DACOROGNA, M. M., DAVÉ, R. R., MÜLLER, U. A., OLSEN, R. B. & PICTET, O. V. (1997) From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and stochastics,* 1**,** 95-129.

GUO, S. *Survival Analysis*, Oxford University Press.

GUPTA, J., VAN BEUKERING, P., VAN ASSELT, H., BRANDER, L., HESS, S. & VAN DER LEEUW, K. (2008) Flexibility mechanisms and sustainable development: lessons from five AIJ projects. *Climate Policy,* 8**,** 261-276.

GUPTA, R. B. & SHIM, J. J. (2007) *Solubility in supercritical carbon dioxide*, CRC.

GWILYM, A. O. & THOMAS, S. (2002) An empirical comparison of quoted and implied bid-ask spreads on futures contracts. *Journal of International Financial Markets, Institutions and Money,* 12**,** 81-99.

HAEFELI, S., BOSI, M. & PHILIBERT, C. (2004) Carbon Dioxide Capture and Storage Issues–Accounting and Baselines under the United Nations Framework Convention on Climate Change (UNFCCC).

HAMAO, Y., MASULIS, R. W. & NG, V. (1990) Correlations in price changes and volatility across international stock markets. *Review of Financial Studies,* 3**,** 281.

HAMILTON, K., BAYON, R., TURNER, G. & HIGGINS, D. (2007) State of the voluntary Carbon Markets 2007. *Picking up steam. Washington, DC*.

HAMILTON, K., SJARDIN, M., MARCELLO, T. & XU, G. (2008) Forging a frontier: State of the voluntary Carbon Markets 2008. *Ecosystem Marketplace & New Carbon Finance*.

HANOUSEK, J. & PODPIERA, J. (2001) Detection of Bank Failures in Transition Economies: The Case of the Czech Republic. *Czech Journal of Economics and Finance (Finance a uver),* 51**,** 264-278.

HANOUSEK, J. & PODPIERA, R. (2004) Czech experience with market maker trading system. *Economic Systems,* 28**,** 177-191.

HANSEN, L. P. (1982) Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society***,** 1029-1054.

HARRIS, L. (1986) A transaction data study of weekly and intradaily patterns in stock returns. *Journal of Financial Economics,* 16**,** 99-117.

HARRIS, L. (1988) Dangers of Regulatory Overreaction to the October 1987 Crash. *Cornell L. Rev.,* 74**,** 927.

HARRIS, L. (1989) The October 1987 S&P 500 stock-futures basis. *The Journal of Finance,* 44**,** 77-99.

HARRIS, L. (1989) S&P 500 cash stock price volatilities. *Journal of Finance,* 44**,** 1155-1175.

HARRIS, L. (1990) Estimation of stock price variances and serial covariances from discrete observations. *Journal of Financial and Quantitative Analysis,* 25**,** 291-306.

HARRIS, L. (1991) Stock price clustering and discreteness. *Review of Financial Studies,* 4**,** 389.

HARRIS, L. (2003) *Trading and exchanges: market microstructure for practitioners*, Oxford University Press.

HARRIS, L. & HASBROUCK, J. (1996) Market vs. limit orders: the SuperDOT evidence on order submission strategy. *Journal of Financial and Quantitative Analysis,* 31**,** 213-231.

HARRIS, L., SOFIANOS, G. & SHAPIRO, J. E. (1994) Program trading and intraday volatility. *Review of Financial Studies,* 7**,** 653.

HARRIS, L. E. (1994) Minimum price variations, discrete bid-ask spreads, and quotation sizes. *Review of Financial Studies,* 7**,** 149.

HARRIS, M. & RAVIV, A. (1988) Corporate control contests and capital structure. *Journal of Financial Economics,* 20**,** 55-86.

HARRIS, M. & RAVIV, A. (1993) Differences of opinion make a horse race. *Review of Financial Studies,* 6**,** 473.

HARRIS, M. S. (1998) The association between competition and managers' business segment reporting decisions. *Journal of Accounting Research,* 36**,** 111-128.

HARVEY, A., RUIZ, E. & SHEPHARD, N. (1994) Multivariate stochastic variance models. *The Review of Economic Studies,* 61**,** 247.

HASBROUCK, J. (1991a) Measuring the information content of stock trades. *Journal of Finance,* 46**,** 179-207.

HASBROUCK, J. (1987) A Note on Forecaster Discord and Consensus Prediction Error. *Journal of Business & Economic Statistics,* 5**,** 151-154.

HASBROUCK, J. (1987) Order arrival, quote behavior, and the return-generating process. *The Journal of Finance,* 42**,** 1035-1048.

HASBROUCK, J. (1991b) The summary informativeness of stock trades: An econometric analysis. *Review of Financial Studies,* 4**,** 571.

HASBROUCK, J. (1988) Trades, quotes, inventories, and information. *Journal of Financial Economics,* 22**,** 229-252.

HASBROUCK, J. (1992) Using the TORQ database. Citeseer.

HASBROUCK, J. (1993) Assessing the quality of a security market: A new approach to transaction-cost measurement. *Review of Financial Studies,* 6**,** 191.

HASBROUCK, J. (1995) One security, many markets: Determining the contributions to price discovery. *The Journal of Finance,* 50**,** 1175-1199.

HASBROUCK, J. (1996) Modeling market microstructure time series. *Handbook of statistics,* 14**,** 647-692.

HASBROUCK, J. (1996) Order characteristics and stock price evolution An application to program trading* 1. *Journal of Financial Economics,* 41**,** 129-149.

HASBROUCK, J. (2007) *Empirical market microstructure,* Oxford University Press.

HASBROUCK, J. (2007) *Empirical market microstructure: The institutions, economics and econometrics of securities trading,* Oxford University Press, USA.

HASBROUCK, J. & HO, T. S. Y. (1987) Order arrival, quote behavior, and the return-generating process. *Journal of Finance,* 42**,** 1035-1048.

HASBROUCK, J. & HO, T. S. Y. (1987) ZOrder arrival, qoute behavior, and the return0generating process+. *The Journal of Finance,* 42.

HASBROUCK, J. & SAAR, G. (2002) Limit orders and volatility in a hybrid market: The Island ECN. *unpublished paper, New York University*.

HASBROUCK, J. & SOFIANOS, G. (1993) The trades of market makers: An empirical analysis of NYSE specialists. *Journal of Finance***,** 1565-1593.

HAUSMAN, J. A., LO, A. W. & MACKINLAY, A. C. (1992) An ordered probit analysis of transaction stock prices. *Journal of financial economics,* 31**,** 319-379.

HAUTSCH, N. (2000) Modelling intraday trading activity using Box-Cox ACD models.

HAUTSCH, N. (2002) *Modelling intraday trading activity using Box-Cox-ACD models,* Universität Konstanz, Fachbereich für Wirtschaftswissenschaften.

HAUTSCH, N. (2004) *Modelling irregularly spaced financial data: theory and practice of dynamic duration models*, Springer Verlag.

HAUTSCH, N. & HESS, D. (2002) The processing of non-anticipated information in financial markets: Analyzing the impact of surprises in the employment report. *European Finance Review,* 6**,** 133.

HAUTSCH, N. & POHLMEIER, W. (2002) Econometric Analysis of Financial Transaction Data: Pitfalls and Opportunities. Allg. Stat.

HAYEK, F. A. (1945) The use of knowledge in society. *The American Economic Review***,** 519-530.

HAYEK, F. A. (1945) von (1945). The use of knowledge in society. *American Economic Review,* 35**,** 519-530.

HEDVALL, K., NIEMEYER, J. & ROSENQVIST, G. (1997) Do buyers and sellers behave similarly in a limit order book? A high-frequency data examination of the Finnish stock exchange. *Journal of Empirical Finance,* 4**,** 279-293.

HILLIER, B. (1997) *The economics of asymmetric information*, Macmillan.

HIRSHLEIFER, J. & RILEY, J. G. (1992) *The analytics of uncertainty and information,* Cambridge University Press.

HO, S. P. S. (1981) South Korea and Taiwan: Development prospects and problems in the 1980s. *Asian Survey,* 21**,** 1175-1196.

HO, T. & STOLL, H. (1981) VOptimal Dealer Pricing Under Transactions and Return UncertaintyV. *Journal of Financial Economics,* 9**,** 47.

HO, T., STOLL, H. R. & RODNEY, L. W. C. F. F. R. (1979) *Optimal dealer pricing under transactions and return uncertainty*, Salomon Brothers Center for the Study of Financial Institutions.

HO, T. S. Y. & MACRIS, R. G. (1984) Dealer bid-ask quotes and transaction prices: An empirical study of some AMEX options. *Journal of Finance***,** 23-45.

HO, T. S. Y. & SAUNDERS, A. (1981) The determinants of bank interest margins: theory and empirical evidence. *Journal of Financial and Quantitative Analysis,* 16**,** 581-600.

HO, T. S. Y. & SAUNDERS, A. (1985) A micro model of the federal funds market. *Journal of Finance,* 40**,** 977-988.

HO, T. S. Y., SCHWARTZ, R. A. & WHITCOMB, D. K. (1985) The trading decision and market clearing under transaction price uncertainty. *Journal of Finance***,** 21-42.

HO, T. S. Y. & STOLL, H. R. (1983) The dynamics of dealer markets under competition. *Journal of Finance,* 38**,** 1053-1074.

HOGAN, K. C. (1994) Sources of meteor showers and heat waves in the foreign exchange market. *Journal of International Economics,* 37**,** 239-247.

HOLDEN, C. W. & SUBRAHMANYAM, A. (1992) Long-lived private information and imperfect competition. *Journal of Finance***,** 247-270.

HOLDEN, C. W. & SUBRAHMANYAM, A. (1994) Risk aversion, imperfect competition, and long-lived information. *Economics Letters,* 44**,** 181–190.

HOLDER, M. E., QI, M. & SINHA, A. K. (2004) The Impact of Time Duration between trades on the price of Treasury Note Futures contracts. *Journal of Futures Markets,* 24**,** 965-980.

HOLTHAUSEN, R. W., LEFTWICH, R. W. & MAYERS, D. (1987) The effect of large block transactions on security prices: A cross-sectional analysis. *Journal of Financial Economics,* 19**,** 237-267.

HUANG, R. D. & STOLL, H. R. (1994) Market microstructure and stock return predictions. *Review of Financial Studies,* 7**,** 179.

HUANG, R. D. & STOLL, H. R. (1996) Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE* 1. *Journal of Financial Economics,* 41**,** 313-357.

HUANG, R. D. & STOLL, H. R. (1997) The components of the bid-ask spread: A general approach. *Review of Financial Studies,* 10**,** 995.

HUBER, C., LIMNIOS, N. & MESBAH, M. (2008) *Mathematical methods in survival analysis, reliability and quality of life*, ISTE.

HUJER, R., BLIEN, U., CALIENDO, M. & ZEISS, C. (2002) Macroeconometric Evaluation of Active Labour Market Policies in Germany–.

HUJER, R. & VULETIC, S. (2005) The regime-switching ACD framework: The use of the comprehensive family of distributions. University of Frankfurt Working Paper.

HUJER, R. & VULETIC, S. (2007) Econometric analysis of financial trade processes by discrete mixture duration models. *Journal of Economic Dynamics and Control,* 31**,** 635-667.

IOSIFESCU, M., LIMNIOS, N. & OPRIÈ™AN, G. (2009) *Introduction to Stochastic Models*, ISTE.

ISENEGGER, P. & VON WYSS, R. (2009) The Valuation of Derivatives on Carbon Emission Certificates-a GARCH Approach.

ITO, T., ENGLE, R. F. & LIN, W. L. (1992) Where does the meteor shower come from?:: The role of stochastic policy coordination. *Journal of International Economics,* 32**,** 221-240.

ITO, T., LYONS, R. K. & MELVIN, M. T. (1998) Is there private information in the FX market? The Tokyo experiment. *The Journal of Finance,* 53**,** 1111-1130.

JACKSON, M. O. (1991) Bayesian implementation. *Econometrica: Journal of the Econometric Society,* 59**,** 461-477.

JACQUIER, E., POLSON, N. G. & ROSSI, P. E. (1994) [Bayesian Analysis of Stochastic Volatility Models]: Reply. *Journal of Business & Economic Statistics,* 12**,** 413-417.

JAIN, P. C. & JOH, G. H. (1988) The dependence between hourly prices and trading volume. *Journal of Financial and Quantitative Analysis,* 23**,** 269-283.

JANSSEN, J. (1986) *Semi-Markov models: theory and applications*, Plenum Press.

JANSSEN, J. & LIMNIOS, N. (1999) *Semi-Markov models and applications*, Kluwer Academic Publishers.

JANSSEN, J. & MANCA, R. (2006) *Applied semi-Markov processes*, Springer Science+Business Media.

JANSSEN, J. & MANCA, R. (2007) *Semi-Markov risk models for finance, insurance and reliability*, Springer.

JANSSEN, J., MANCA, R., VOLPE, E. & PRIGNANO, E. V. (2009) *Mathematical Finance*, ISTE.

JASIAK, J. (1998) Persistence in intertrade durations. *FINANCE-PARIS-,* 19**,** 165-196.

JAYNES, E. T. (1994) Probability theory: The logic of Science http://omega. albany. edu: 8008. *JaynesBook. html*.

JEWELL, N. P. (1996) *Lifetime data: models in reliability and survival analysis*, Kluwer Academic.

JOHANSEN, S. (1988) Statistical analysis of cointegration vectors. *Journal of economic dynamics and control,* 12**,** 231-254.

JOHNSON, N. L., KOTZ, S. & BALAKRISHNAN, N. Continuous univariate distributions (Volume 1,). IMA.

JONES, C. M., KAUL, G. & LIPSON, M. L. (1994) Transactions, volume, and volatility. *Review of Financial Studies,* 7**,** 631.

JONG, F. & RINDI, B. (2009) The Microstructure of Financial Markets.

JONG, F. & RINDI, B. (2009) *The Microstructure of Financial Markets*, Cambridge University Press.

JORION, P. (1996) Risk and turnover in the foreign exchange market. University of Chicago Press.

JUN, B. & VIVES, X. (2004) Strategic Incentives in Dynamic Oligopoly". *Journal of Economic Theory,* 116**,** 249–281.

KALBFLEISCH, J. D. & PRENTICE, R. L. (2002) *The statistical analysis of failure time data*, J. Wiley.

KAMARA, A. (1998) Time Variation in the Turn-of-the-Year Seasonal, University of Washington and University of California. Los Angeles, Mimeo.

KARA, M., SYRI, S., LEHTILÄ, A., HELYNEN, S., KEKKONEN, V., RUSKA, M. & FORSSTRÖM, J. (2008) The impacts of EU CO2 emissions trading on electricity markets and electricity consumers in Finland. *Energy Economics,* 30**,** 193-211.

KARATZAS, I. & SHREVE, S. E. (1991) *Brownian motion and stochastic calculus*, Springer.

KASWAN, A. (2010) Decentralizing Cap-and-Trade? State Controls within a Federal

Greenhouse Gas Cap-and-Trade Program. *Virginia Environmental Law Journal,* 28.

KAWALLER, I. G., KOCH, P. D. & KOCH, T. W. (1987) The temporal price relationship between S&P 500 futures and the S&P 500 index. *The Journal of Finance,* 42**,** 1309-1329.

KAWALLERPAUL, D. & IRA, G. (1990) Intraday relationships between volatility in S&P 500 futures prices and volatility in the S&P 500 index* 1. *Journal of Banking & Finance,* 14**,** 373-397.

KAYAHAN, B., STENGOS, T. & SALTO LU, B. (2002) Intra-day features of realized volatility: evidence from an emerging market. *International Journal of Business,* 1**,** 17-24.

KE, M. C., JIANG, C. H. & HUANG, Y. S. (2004) The impact of tick size on intraday stock price behavior: evidence from the Taiwan Stock Exchange. *Pacific-Basin Finance Journal,* 12**,** 19-39.

KEIM, D. B. (1983) Size-related anomalies and stock return seasonality:: Further empirical evidence. *Journal of Financial Economics,* 12**,** 13-32.

KEIM, D. B. & MADHAVAN, A. (1996) The upstairs market for large-block transactions: Analysis and measurement of price effects. *Review of Financial Studies,* 9**,** 1.

KEIM, D. B. & STAMBAUGH, R. F. (1984) A further investigation of the weekend effect in stock returns. *Journal of Finance,* 39**,** 819-835.

KENDALL, M. G. & GIBBONS, J. D. (1962) Rank correlation methods. Griffin London.

KEPPLER, J. H. & MANSANET-BATALLER, M. Causalities between CO2, electricity, and other energy variables during Phase I and Phase II of the EU ETS. *Energy Policy,* 38**,** 3329-3341.

KESTNER, L. (2003) *Quantitative trading strategies: harnessing the power of quantitative techniques to create a winning trading program,* McGraw-Hill.

KIM, J., ALLENBY, G. M. & ROSSI, P. E. (2002) Modeling consumer demand for variety. *Marketing Science,* 21**,** 229-250.

KIM, S. H. & OGDEN, J. P. (1996) Determinants of the components of bid ask spreads on stocks. *European Financial Management,* 2**,** 127-145.

KIM, W. & WEI, S. J. (2002) Foreign portfolio investors before and during a crisis. *Journal of International Economics,* 56**,** 77-96.

KISSELL, R., GLANTZ, M. & MALAMUT, R. (2003) *Optimal trading strategies: quantitative approaches for managing market impact and trading risk,* AMACOM.

KLEIDON, A. & WERNER, I. (1996) UK and US trading of British cross-listed stocks: An intraday analysis of market integration. *Review of Financial Studies,* 9**,** 619-664.

KOSOBUD, R., STOKES, H. & SCOTT, B. (2005) Valuing Tradable Private Rights to Pollute the Public's Air. *Review of Accounting and Finance,* 4.

KRUGER, J., OATES, W. E. & PIZER, W. A. (2007) Decentralization in the EU emissions trading scheme and lessons for global policy. *Review of Environmental Economics and Policy,* 1**,** 112.

KUMAR, P. & SEPPI, D. J. (1994) Information and index arbitrage. *Journal of Business,* 67**,** 481-509.

KWIATKOWSKI PETER, C. B. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root* 1:: How sure are we that economic time series have a unit root? *Journal of econometrics,* 54**,** 159-178.

KYLE, A. S. (1984) Market structure, information, futures markets, and price formation. *International agricultural trade: Advanced readings in price formation, market structure, and price instability***,** 45-64.

KYLE, A. S. (1985) Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society,* 53**,** 1315-1335.

KYLE, A. S. (1989) Informed speculation with imperfect competition. *The Review of Economic Studies*, 56**,** 317-355.

KYLE, A. S. & VILA, J. L. (1991) Noise trading and takeovers. *The RAND Journal of Economics,* 22**,** 54-71.

KYLE, A. S. & VISWANATHAN, S. (2008) How to define illegal price manipulation. *The American Economic Review,* 98**,** 274-279.

LAKONISHOK, J. & LEVI, M. (1982) Weekend effects on stock returns: a note. *The Journal of Finance,* 37**,** 883-889.

LAKONISHOK, J. & MABERLY, E. (1990) The weekend effect: Trading patterns of individual and institutional investors. *The Journal of Finance,* 45**,** 231-243.

LAKONISHOK, J. & SMIDT, S. (1984) Volume and turn-of-the-year behavior. *Journal of Financial Economics,* 13**,** 435-455.

LAKONISHOK, J. & SMIDT, S. (1988) Are seasonal anomalies real? A ninety-year perspective. *Review of Financial Studies,* 1**,** 403.

LAMOUREUX, C. G. & LASTRAPES, W. D. (1990) Heteroskedasticity in stock return data: volume versus GARCH effects. *The Journal of Finance,* 45**,** 221-229.

LANCASTER, T. (1997) The econometric analysis of duration data. Cambridge University Press.

LANDSIEDEL, C. (2005) The Possible Accession of Turkey to the EU-A Turkish and a German Perspective.

LAUX, P. A. & NG, L. K. (1993) The sources of GARCH: Empirical evidence from an intraday returns model incorporating systematic and unique risks. *Journal of International Money and Finance,* 12**,** 543-560.

LEACH, J. C. & MADHAVAN, A. N. (1993) Price experimentation and security market structure. *Review of Financial Studies,* 6**,** 375.

LEBARON, B. D. & UNIVERSITY OF WISCONSIN--MADISON. SOCIAL SYSTEMS RESEARCH, I. (1992) *Do Moving Average Trading Rule Results Imply Nonlinearites in Foreign Exchange Markets?*, Citeseer.

LECOCQ, F. (2005) *State and Trends of the Carbon Market 2004*, World Bank Publications.

LECOCQ, F. & AMBROSI, P. (2007) Policy Monitor Edited by Maureen Cropper The Clean Development Mechanism: History, Status, and Prospects.

LECOCQ, F. & CAPOOR, K. (2005) State and Trends of the Carbon Market 2005.

LEE, C., SHLEIFER, A. & THALER, R. H. (1990) Investor sentiment and the closed-end fund puzzle. National Bureau of Economic Research Cambridge, Mass., USA.

LIESENFELD, R. & RICHARD, J. F. (2003) Univariate and multivariate stochastic volatility models: estimation and diagnostics. *Journal of empirical finance,* 10**,** 505-531.

LIN, J. C., SANGER, G. C. & BOOTH, G. G. (1995) Trade size and components of the bid-ask spread. *Review of Financial Studies,* 8**,** 1153.

LIN, W. L., ENGLE, R. F. & ITO, T. (1994) Do bulls and bears move across borders? International transmission of stock returns and volatility. *Review of Financial Studies,* 7**,** 507.

LINS, K., STRICKLAND, D. & ZENNER, M. (2005) Do non-US firms issue stock on US equity markets to relax capital constraints. *Journal of Financial and Quantitative Analysis,* 40**,** 109-133.

LIU, H. X., RECKER, W. & CHEN, A. (2004) Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data. *Transportation Research Part A: Policy and Practice,* 38**,** 435-453.

LJUNG, G. M. & BOX, G. E. P. (1978) On a measure of lack of fit in time series models. Biometrika Trust.

LO, A. W. & MACKINLAY, A. C. (1988) Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of financial studies,* 1**,** 41.

LUCA, G., RIVIECCIO, G. & ZUCCOLOTTO, P. (2008) Exploring the Copula Approach for the Analysis of Financial Durations. *Mathematical and Statistical Methods in Insurance and Finance***,** 99-106.

LUNDE, A. (2000) A Generalized Gamma Autoregressive Conditional Duration Model Discussion Paper [Z]. *Aarlborg University*.

LUSCH, R. F. & TAY, N. S. P. (2004) Agent-based modeling: Gaining insight into firm and industry performance. *Assessing marketing strategy performance***,** 212–27.

LUU, J. C. & MARTENS, M. (2003) Testing the mixture of distributions hypothesis using “realized” volatility. *Journal of Futures Markets,* 23**,** 661-679.

LYONS, R. K. (1993) Optimal transparency in a dealership market with an application to foreign exchange. NBER.

LYONS, R. K. (1995) Tests of microstructural hypotheses in the foreign exchange market. *Journal of Financial Economics,* 39**,** 321-351.

LYONS, R. K. (1996) Foreign exchange volume: sound and fury signifying nothing? , University of Chicago Press.

LYONS, R. K. (1996) Optimal Transparency in a Dealer Market with an Application to Foreign Exchange* 1. *Journal of Financial Intermediation,* 5**,** 225-254.

MACHIN, D., CHEUNG, Y. B. & PARMAR, M. K. B. (2006) *Survival analysis: a practical approach*, Wiley.

MADDALA, G. S. & RAO, C. R. GARCH Models of Volatility.

MADHAVAN, A. (1992) Trading mechanisms in securities markets. *Journal of Finance,* 47**,** 607-641.

MADHAVAN, A. (1995) Consolidation, fragmentation, and the disclosure of trading information. *Review of Financial Studies,* 8**,** 579.

MADHAVAN, A. (2000) Market microstructure: A survey* 1. *Journal of Financial Markets,* 3**,** 205-258.

MADHAVAN, A. & CHENG, M. (1997) In search of liquidity: Block trades in the upstairs and downstairs markets. *Review of Financial Studies,* 10**,** 175.

MADHAVAN, A. & PANCHAPAGESAN, V. (2000) Price discovery in auction markets: A look inside the black box. *Review of Financial Studies,* 13**,** 627.

MADHAVAN, A., RICHARDSON, M. & ROOMANS, M. (1997) Why do security prices change? A transaction-level analysis of NYSE stocks. *Review of Financial Studies,* 10**,** 1035.

MADHAVAN, A. & SMIDT, S. (1991) A Bayesian model of intraday specialist pricing* 1. *Journal of Financial Economics,* 30**,** 99-134.

MADHAVAN, A. & SMIDT, S. (1993) An analysis of changes in specialist inventories and quotations. *Journal of Finance,* 48**,** 1595-1628.

MADHAVAN, A., TREYNOR, J. L. & WAGNER, W. H. (2007) Execution of portfolio decisions. *Managing Investment Portfolios: A Dynamic Process. John Wiley& Sons, Hoboken, New Jersey***,** 637-681.

MAKRIDAKIS, S. (1993) Accuracy measures: theoretical and practical concerns. Elsevier.

MALLER, R. A., MÜLLER, G. & SZIMAYER, A. (2008) GARCH modelling in continuous time for irregularly spaced time series data. *Bernoulli,* 14**,** 519-542.

MANDELBROT, B. (1963) The variation of certain speculative prices. *The journal of business,* 36**,** 394-419.

MANGANELLI, S. (2005) Duration, volume and volatility impact of trades. *Journal of Financial markets,* 8**,** 377-399.

MANGANELLI, S. (2005) Duration, volume, and the price impact of trades. *Journal of Financial Markets,* 8**,** 377-399.

MANSANET-BATALLER, M., CHEVALLIER, J., HERVE-MIGNUCCI, M. & ALBEROLA, E. (2010) EUA and sCER Phase II Price Drivers: Unveiling the reasons for the existence of the EUA-sCER spread. *Energy Policy*.

MANSANET-BATALLER, M. & PARDO, A. (2007) The Effects of National Allocation Plans on Carbon-Markets. *Unpublished manuscript. papers. ssrn. com/sol3/papers. cfm*.

MANSANET-BATALLER, M. & PARDO, A. (2008) CO2 prices and portfolio management. *Unpublished manuscript. http://papers. ssrn. com/sol3/papers. cfm*.

MANSANET-BATALLER, M. & PARDO, A. (2008) What you should know about Carbon Markets. *Energies,* 1**,** 120-153.

MANSANET-BATALLER, M., PARDO, A. & VALOR, E. (2007) CO2 prices, energy and weather. *The Energy Journal,* 28**,** 73-92.

MANSANET-BATALLER, M. & PARDO, T. (2009) Impact of Regulatory Announcements on CO2 prices. *Journal of Energy Markets,* 2**,** 1–33.

MARENGO, L. & TORDJMAN, H. (1996) Speculation, heterogeneity and learning: a simulation model of exchange rates dynamics. *Kyklos,* 49**,** 407-438.

MARIANO, R. S. & TSE, Y. K. (2008) *Econometric forecasting and high-frequency data analysis*, World Scientific.

MARTENS, M. (1998) Price discovery in high and low volatility periods: open outcry versus electronic trading. *Journal of International Financial Markets, Institutions and Money,* 8**,** 243-260.

MEDDAHI, N., RENAULT, E. & WERKER, B. (2006) GARCH and irregularly spaced data. *Economics Letters,* 90**,** 200-204.

MEITZ, M. (2006) A necessary and sufficient condition for the strict stationarity of a family of GARCH processes. *Econometric Theory,* 22**,** 985-988.

MEITZ, M. & SAIKKONEN, P. (2008) Ergodicity, mixing, and existence of moments of a class of Markov models with applications to GARCH and ACD models. Cambridge Univ Press.

MEITZ, M. & TERÄSVIRTA, T. (2006) Evaluating models of autoregressive conditional duration. *Journal of Business and Economic statistics,* 24**,** 104-124.

MENDELSON, H. (1982) Market behavior in a clearing house. *Econometrica: Journal of the Econometric Society***,** 1505-1524.

MENDELSON, H. (1985) Random competitive exchange: Price distributions and gains from trade. *Journal of Economic Theory,* 37**,** 254-280.

MENDELSON, H. (1987) Consolidation, fragmentation, and market performance. *Journal of Financial and Quantitative Analysis,* 22**,** 189-207.

MENYAH, K. & PAUDYAL, K. (2000) The components of bid-ask spreads on the London Stock Exchange. *Journal of Banking & Finance,* 24**,** 1767-1785.

MILUNOVICH, G. & JOYEUX, R. (2007) Market efficiency and price discovery in the EU carbon futures market. *Unpublished manuscript. http://ssrn. com/abstractq989272*.

MILUNOVICH, G. & JOYEUX, R. (2007) Testing market efficiency and price discovery in European Carbon Markets. *Research Papers*.

MITCHELL, M. L. & MULHERIN, J. H. (1994) The impact of public information on the stock market. American Finance Association.

MOLHO, I. (1997) *The economics of information: lying and cheating in markets and organizations*, Blackwell.

MULLER, U. A., DACOROGNA, M. M., OLSEN, R. B., PICTET, O. V., SCHWARZ, M. & MORGENEGG, C. (1990) Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis* 1. *Journal of Banking & Finance,* 14**,** 1189-1208.

MULLER, U. A. & SGIER, R. (1992) Statistical analysis of intra-day bid-ask spreads in the foreign exchange market. *Unpublished manuscript, Olsen & Associates, Zürich*.

MURTHY, D. N. P., XIE, M. & JIANG, R. (2004) *Weibull models,* J. Wiley.

NELSEN, R. B. (1999) An Introduction to Copulas Springer. *New York*.

NELSON, D. B. (1991) Conditional heteroskedasticity in asset returns: a new approach. *Econometrica: Journal of the Econometric Society,* 59**,** 347-370.

NEWEY, W. K. & WEST, K. D. (1987) Hypothesis testing with efficient method of moments estimation. *International Economic Review,* 28**,** 777-787.

NING, Q. (2004) Estimation of the stochastic conditional duration model via alternative methods-ECF and GMM. mimeo.

NORDÉN, L. (2003) Asymmetric option price distribution and bid-ask quotes: consequences for implied volatility smiles* 1. *Journal of Multinational Financial Management,* 13**,** 423-441.

O'HARA, M. (1995) *Market microstructure theory*, Wiley.

O'HARA, M. (2003) Presidential address: Liquidity and price discovery. *The Journal of Finance,* 58**,** 1335-1354.

O'HARA, M. (2007) *Market microstructure theory*, Blackwell.

O'HARA, M. & OLDFIELD, G. S. (1986) The microeconomics of market making. *Journal of Financial and Quantitative Analysis,* 21**,** 361-376.

ORPHANIDES, A. & NORDEN, S. (2002) The unreliability of output-gap estimates in real time. *Review of economics and statistics,* 84**,** 569-583.

PACURAR, M. (2008) Autoregressive conditional duration models in finance: a survey of the theoretical and empirical literature. *Journal of Economic Surveys,* 22**,** 711-751.

PAGANO, M. & ROELL, A. (1992) Auction and dealership markets: what is the difference? *European Economic Review,* 36**,** 613-623.

PAGANO, M. & RÖELL, A. (1996) Transparency and liquidity: A comparison of auction and dealer markets with informed trading. *The Journal of Finance,* 51**,** 579-611.

PAGANO, M. & RÖELL, A. A. (1991) Dually-traded Italian equities: London vs. Milan. *CEPR Discussion Papers*.

PAI, J. S., POLASEK, P. & KOZUMI, H. (1995) Irregularly Spaced AR and ARCH (ISAR-ARCH) Models.

PAOLELLA, M. S. & TASCHINI, L. (2008) An econometric analysis of emission allowance prices. *Journal of Banking & Finance,* 32**,** 2022-2032.

PARLOUR, C. A. (1998) Price dynamics in limit order markets. *Review of Financial Studies,* 11**,** 789.

PASCUAL, R., ESCRIBANO, Á. & TAPIA, M. (2004) On the bi-dimensionality of liquidity. *The European Journal of Finance*, 10**,** 542-566.

PESARAN, B. & ROBINSON, G. (1993) The European exchange rate mechanism and the volatility of the sterling-deutschemark exchange rate. *The Economic Journal,* 103**,** 1418-1431.

PHLIPS, L. (1988) *The economics of imperfect information*, Cambridge University Press.

PINDER, S. (2003) An empirical examination of the impact of market microstructure changes on the determinants of option bid-ask spreads. *International Review of Financial Analysis,* 12**,** 563-577.

POJARLIEV, M. & POLASEK, W. (1995) Value at risk estimation for stock indices using the basle committee proposal from.

POLE, A. (2007) *Statistical arbitrage: algorithmic trading insights and techniques*, John Wiley & Sons.

PORTER, D. C. (1992) The probability of a trade at the ask: An examination of interday and intraday behavior. *Journal of Financial and Quantitative Analysis,* 27**,** 209-227.

RACICOT, F. É., THÉORET, R. & COËN, A. (2008) Forecasting irregularly spaced UHF financial data: Realized volatility vs UHF-GARCH models. *International Advances in Economic Research,* 14**,** 112-124.

REISS, P. C. & WERNER, I. M. (1998) Does risk sharing motivate interdealer trading? *Journal of Finance,* 53**,** 1657-1703.

RENAULT, E. & WERKER, B. J. M. (2002) Stochastic volatility with transaction time risk. Discussion paper, University of Montreal.

RENAULT, E. & WERKER, B. J. M. (2005) Stochastic volatility models with transaction time risk. *Department of Economics, University of North Carolina*.

REZEK, J. (1999) Shadow Prices of Sulfur Dioxide Allowances in Phase I Electric Utilities

Annual Meeting of the American Agricultural Economics Association.

RINNE, H. (2009) *The Weibull distribution: a handbook*, CRC Press.

RITTLER, D. (2009) Price Discovery, Causality and Volatility Spillovers in European Union Allowances Phase II: A High Frequency Analysis. *University of Heidelberg, Department of Economics, Discussion Paper Series,* 492.

ROCHET, J. C. & VILA, J. L. (1994) Insider trading without normality. *The Review of Economic Studies,* 61**,** 131-152.

ROLL, R. (1983) Vas ist das? *The Journal of Portfolio Management,* 9**,** 18-28.

ROLL, R. (1984) A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance,* 39**,** 1127-1139.

ROSS, S. M. (2000) *Introduction to probability models*, Harcourt/Academic Press.

ROTFUß, W. (2009) *Intraday price formation and volatility in the European Union emissions trading scheme: an introductory analysis*, Zentrum für Europ. Wirtschaftsforschung.

ROTFUß, W., CONRAD, C. & RITTLER, D. (2009) The European Commission and EUA Prices: A High-Frequency Analysis of the EC's Decision on Second National Allocation Plans.

RUBIN, J. D. (1996) A model of intertemporal emission trading, banking, and borrowing. *Journal of Environmental Economics and Management,* 31**,** 269-286.

RUSSELL, J. R. & ENGLE, R. F. (2005) A Discrete-State Continuous-Time Model of Financial Transactions Prices and Times. *Journal of Business and Economic Statistics,* 23**,** 166-180.

RUSSELL, J. R. & ENGLE, R. F. (2005) A discrete-state continuous-time model of financial transactions prices and times: the ACM-ACD model. *Journal of Business and Economic Statistics,* 23**,** 166–180.

RUSSELL, J. R., ENGLE, R. F., BANDI, F. & RUSSELL, J. (1998) Econometric analysis of discrete-valued irregularly-spaced financial transactions data using a new autoregressive conditional multinomial model.

RUST, J. (1994) Estimation of dynamic structural models, problems and prospects: discrete decision processes.

SANDMANN, G. & KOOPMAN, S. J. (1998) Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics,* 87**,** 271-301.

SANDRONI, A. (2005) Market selection when markets are incomplete. *Journal of Mathematical Economics,* 41**,** 91-104.

SANIN, M. E. & VIOLANTE, F. (2009) Understanding volatility dynamics in the EU-ETS market: lessons from the future. CORE Discussion Paper.

SARIN, A., SHASTRI, K. A. & SHASTRI, K. (2000) Ownership structure and stock market liquidity. *unpublished paper, University of Pittsburgh*.

SAVEDOFF, W. D. & SCHULTZ, T. P. (2000) *Wealth from health: linking social investments to earnings in Latin America*, Inter-Amer Development Bank.

SCHENNACH, S. M. (2000) The economics of pollution permit banking in the context of Title IV of the 1990 Clean Air Act Amendments. *Journal of Environmental Economics and Management,* 40**,** 189-210.

SCHLEICH, J., EHRHART, K. M., HOPPE, C. & SEIFERT, S. (2006) Banning banking in EU emissions trading? *Energy Policy,* 34**,** 112-120.

SCHWARZ, T. V. & SZAKMARY, A. C. (1994) Price discovery in petroleum markets: arbitrage, cointegration, and the time interval of analysis. *Journal of Futures Markets,* 14**,** 147-167.

SEIFERT, J., UHRIG-HOMBURG, M. & WAGNER, M. (2008) Dynamic behavior of CO2 spot prices. *Journal of Environmental Economics and Management,* 56**,** 180-194.

SEPPI, D. J. (1990) Equilibrium block trading and asymmetric information. *the Journal of Finance,* 45**,** 73-94.

SEPPI, D. J. (1992) Block trading and information revelation around quarterly earnings announcements. *Review of Financial Studies,* 5**,** 281.

SEREDNYAKOV, A. (2005) A model of the components of the bid-ask spread. Working paper, University of Minnesota.

SHILLER, R. (1984) Stock Market and Social Dynamics. *Brookings Papers on Ecomomic Activity (2)***,** 457-498.

SHILLER, R. J. (1981) Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review,* 71**,** 421-436.

SHILLER, R. J. (1984) Theories of aggregate stock price movements. *The Journal of Portfolio Management,* 10**,** 28-37.

SILVA, A. C. & CHAVEZ, G. (2002) Components of execution costs: evidence of asymmetric information at the Mexican Stock Exchange* 1. *Journal of International Financial Markets, Institutions and Money,* 12**,** 253-278.

SMIDT, S. (1971) Which road to an efficient stock market: free competition or regulated monopoly? *Financial Analysts Journal,* 27**,** 18-69.

SMITH, L. & SØRENSEN, P. (2000) Pathological outcomes of observational learning. *Econometrica,* 68**,** 371-398.

SOFIANOS, G. & WERNER, I. M. (2000) The trades of NYSE floor brokers* 1. *Journal of Financial Markets,* 3**,** 139-176.

SPIEGEL, M. & SUBRAHMANYAM, A. (1992) Informed speculation and hedging in a noncompetitive securities market. *Review of Financial Studies,* 5**,** 307.

SPIEGEL, M. & SUBRAHMANYAM, A. (1995) On intraday risk premia. *The Journal of Finance,* 50**,** 319-339.

SPIERDIJK, L. (2004) An empirical analysis of the role of the trading intensity in information dissemination on the NYSE. *Journal of empirical finance,* 11**,** 163-184.

SPULBER, D. F. (1999) *Market microstructure: intermediaries and the theory of the firm*, Cambridge University Press.

STIGLER, G. J. (1967) Imperfections in the capital market. *The Journal of Political Economy,* 75**,** 287-292.

STOCK, J. H. (1988) Estimating continuous-time processes subject to time deformation: an application to postwar US GNP. American Statistical Association.

STOCK, J. H. & WATSON, M. W. (1988) Testing for common trends. *Journal of the American statistical Association,* 83**,** 1097-1107.

STOLL, H. R. (1978) The pricing of security dealer services: An empirical study of NASDAQ stocks. *Journal of Finance,* 33**,** 1153-1172.

STOLL, H. R. (1989) Inferring the components of the bid-ask spread: Theory and empirical tests. *Journal of Finance,* 44**,** 115-134.

STOLL, H. R. (1992) Principles of trading market structure. *Journal of financial services research,* 6**,** 75-106.

STOLL, H. R. & WHALEY, R. E. (1990) Stock market structure and volatility. *Review of Financial Studies*, 3**,** 37.

STRICKLAND, C. M., FORBES, C. S. & MARTIN, G. M. (2006) Bayesian analysis of the stochastic conditional duration model. *Computational statistics & data analysis,* 50**,** 2247-2267.

STRICKLAND, L. A. (2005) *Out of the Cubicle and Into Business: 114 Questions to Answer Before You Make the Move from a Corporation Or University Job Into Your Own Business!*, Authorhouse.

SUBRAHMANYAM, A. (1991a) Risk aversion, market liquidity, and price efficiency. *Review of Financial Studies,* 4**,** 417.

SUBRAHMANYAM, A. (1991b) A theory of trading in stock index futures. *Review of Financial Studies,* 4**,** 17.

SULLIVAN, R., TIMMERMANN, A., WHITE, H. & UNIVERSITY OF CALIFORNIA, S. D. D. O. E. (1998) The dangers of data-driven inference: the case of calendar effects in stock returns. *LSE FINANCIAL MARKETS GROUP DISCUSSION PAPER SERIES*.

TAIYAB, N. (2006) *Exploring the market for voluntary carbon offsets*, Iied.

TAUCHEN, G. E. & PITTS, M. (1983) The price variability-volume relationship on speculative markets. *Econometrica: Journal of the Econometric Society***,** 485-505.

TAY, A. (2004) Nonparametric Conditional Density Estimation, with an Application to Stock Returns Volatility.

TAY, A., TING, C., TSE, Y. K. & WARACHKA, M. (2009) Using high-frequency transaction data to estimate the probability of informed trading. Oxford Univ Press.

TAYLOR, H. M. & KARLIN, S. (1994) *An introduction to stochastic modeling*, Academic Press.

THALER, R. (1987) Anomalies: seasonal movements in security prices ii: weekend, holiday, turn of the month, and intraday effects. *The Journal of Economic Perspectives,* 1**,** 169-177.

THALER, R. H. (1987) Amomalies: The january effect. *The Journal of Economic Perspectives,* 1**,** 197-201.

THALER, R. H. (1994) *Quasi rational economics*, Russell Sage Foundation Publications.

THEISSEN, E. (2002) Price discovery in floor and screen trading systems. *Journal of Empirical Finance,* 9**,** 455-474.

TIJMS, H. C. (2003) *A first course in stochastic models*, Wiley.

TIM, B. & RICHARD, T. (1990) A multivariate generalized ARCH approach to modeling risk premia in forward foreign exchange rate markets. *Journal of International Money and Finance,* 9**,** 309-324.

TINIC, S. M. (1972) The economics of liquidity services. *The Quarterly Journal of Economics,* 86**,** 79-93.

TINIC, S. M. & RICHARD, R. (1972) West, 1972, C ompetition and the pricing of dealer service in the over-the-counter stock ma rket. *Journal of Financial and Quantitative Analysis,* 7**,** 1707-1727.

TINIC, S. M. & WEST, R. R. (1972) Competition and the pricing of dealer service in the over-the-counter stock market. *Journal of Financial and Quantitative Analysis,* 7**,** 1707-1727.

TINIC, S. M. & WEST, R. R. (1974) Marketability of common stocks in Canada and the USA: A comparison of agent versus dealer dominated markets. *Journal of Finance,* 29**,** 729-746.

TOWNSEND, R. M. (1978) Intermediation with costly bilateral exchange. *The Review of Economic Studies,* 45**,** 417-425.

TOWNSEND, R. M. (1978) Market anticipations, rational expectations, and Bayesian analysis. *International Economic Review,* 19**,** 481-494.

TRÜCK, S., BORAK, S., HÄRDLE, W. & WERON, R. (2007) Convenience yields for CO_2 emission allowance futures contracts. *Working Paper, Humboldt-University of Berlin*.

UHRIG-HOMBURG, M. & WAGNER, M. (2006) Success chances and optimal design of derivatives on co2 emission certificates. Working Paper, University of Karlsruhe.

UHRIG-HOMBURG, M. & WAGNER, M. (2007) Futures Price Dynamics of C02 Emission Certificates–An Empirical Analysis. *Woring paper, chair of financial engineering and derivatives, university of karlsruhe*.

UHRIG-HOMBURG, M. & WAGNER, M. (2008) Derivative Instruments in the EU Emissions Trading Scheme–An Early Market Perspective. *Energy & Environment,* 19**,** 635-655.

UHRIG-HOMBURG, M. & WAGNER, M. W. Futures price dynamics of CO2 emission certificates an empirical analysis.

VAN DIJK, D. & FRANSES, P. H. (2002) Smooth transition autoregressive models.

VAN NESS, B. F., VAN NESS, R. A. & WARR, R. S. (2001) How well do adverse-selection components measure adverse-selection? *Financial Management,* 30**,** 77-98.

VAN NORDEN, S. (2005) Are we there yet? Looking for evidence of a New Economy. *Manuscript, HEC Montreal*.

VINOCUR, B. (2009) THE GROUND FLOOR: REITS RALLY, BUT OVER THE LONG HAUL, ARE EXPECTATIONS FOR PROPERTY TOO HIGH? *Barron's,* 80**,** 50.

VINOCUR, J. (2009) ‘German Vote a Turning Point on Russia and Its Energy. *New York Times, March*, 9.

VISWANATHAN, S. (2010) Developing an Efficient Carbon Emissions Allowance Market. *Working Paper Series. Fuqua School of Business Duke University*.

VIVES, X. (1993) How fast do rational agents learn? *The Review of Economic Studies,* 60, 329-347.

VIVES, X. (1993) Information, flexibility, and competition. *Journal of the japanese and International Economies,* 7, 219-219.

VIVES, X. (1996) Social learning and rational expectations. *European Economic Review,* 40, 589-601.

VIVES, X. (1997) Learning from a welfare analysis. *Games Econ. Behav,* 20, 177–200.

VIVES, X. (2001) *Oligopoly pricing: old ideas and new tools,* MIT Press.

VIVES, X. (2005) Complementarities and games: New developments. *Journal of Economic Literature,* 43, 437-479.

VIVES, X. (2008) *Information and Learning in Markets: The Impact of Market Microstructure,* Princeton University Press.

WALD, J. K. & HORRIGAN, H. T. (2005) Optimal limit order choice. *Journal of Business,* 78, 597-620.

WERNER, I. M. & KLEIDON, A. W. (1996) UK and US trading of British cross-listed stocks: An intraday analysis of market integration. *Review of Financial Studies,* 9, 619.

WHITE, H. & DOMOWITZ, I. (1984) Nonlinear regression with dependent observations. *Econometrica: Journal of the Econometric Society*, 143-161.

WING, S. (2008) Explaining the declining energy intensity of the US economy. *Resource and Energy Economics,* 30, 21-49.

WONG, W. K. & CHAN, R. (2008) Markowitz and prospect stochastic dominances. *Annals of Finance,* 4, 105-129.

WONG, W. K., TAN, D. & TIAN, Y. (2009) Informed trading and liquidity in the Shanghai Stock Exchange. Elsevier.

WOOD, R. A., MCINISH, T. H. & ORD, J. K. (1985) An investigation of transactions data for NYSE stocks. *Journal of Finance,* 40, 723-739.

WOOLDRIDGE, J. M. (2001) ASYMPTOTIC PROPERTIES OF WEIGHTED< em> M-ESTIMATORS FOR STANDARD STRATIFIED SAMPLES. *Econometric Theory,* 17, 451-470.

WUENSCHE, O., GRAMMIG, J. & THEISSEN, E. (2007) Time and the Price Impact of a Trade: A Structural Approach. *Discussion Paper*.

WYART, M., BOUCHAUD, J. P., KOCKELKOREN, J., POTTERS, M. & VETTORAZZO, M. (2008) Relation between bid–ask spread, impact and volatility in order-driven markets. *Quantitative Finance,* 8, 41-57.

YAMAGUCHI, K. (1999) *Event history analysis*, Sage Publ.

ZABEL, E. (1981) Competitive price adjustment without market clearing. *Econometrica: Journal of the Econometric Society,* 49, 1201-1221.

ZHANG, A., ZHANG, Y. & ZHAO, R. (2001) Impact of Ownership and Competition on the Productivity of Chinese Enterprises* 1. *Journal of Comparative Economics,* 29, 327-346.

ZHANG, M. Y., RUSSELL, J. R. & TSAY, R. S. (2001) A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics,* 104, 179-207.

ZHANG, Y. J. & WEI, Y. M. An overview of current research on EU ETS: Evidence from its operating mechanism and economic effect. *Applied Energy*, 87**,** 1804-1814.

ZHANG, Z. (2001) Real Exchange Rate Misalignment in China: An Empirical Investigation* 1. *Journal of Comparative Economics*, 29**,** 80-94.

ZUBULAKE, P. & LEE, S. *The High Frequency Game Changer: How Automated Trading Strategies Have Revolutionized the Markets*, John Wiley & Sons.

The Role of Trading Intensity in Duration Modelling and Price Discovery

Evidence from the European Carbon market

Volume II

Appendix

Doctoral Thesis Submitted by

Kalaitzoglou Iordanis Angelos

For the completion of a three years PhD program

July 2011

Heriot-Watt University

School of Management and Languages

Department of Accountancy, Economics and Finance

Supervisors: **Mohamed Sherif**

Boulis Ibrahim

Julian Fennema

Appendix 4.A

Tables

Summary of Models

This section summarizes the models estimated in this chapter (Table 4.1) and the functions $g(\varepsilon_i)$ of the innovations, employed to compute the long-term forecasts of non-linear models in Eq. (4.26) (Table 4.2).

Table 4.1: Summary of Estimated Models

ACD	E W G	$\psi_i = \omega + a\, x_{i-1} + \beta\, \psi_{i-1}$
Log-ACD	E W G	$ln\psi_i = \omega + a\, ln\, x_{i-1} + \beta\, ln\, \psi_{i-1}$
EXACD	E W G	$ln\psi_i = \omega + a\, \varepsilon_{i-1} + \zeta\lvert\varepsilon_{i-1} - 1\rvert + \beta\, ln\, \psi_{i-1}$
ACD-OTC	E W G	$\psi_i = \omega + (a\ + \zeta\, * D_{i-1})x_{i-1} + \beta\, \psi_{i-1}$
BCACD-OTC	E W G	$ln\psi_i = \omega + (a\ + \zeta\, * D_{i-1}) * (\varepsilon_{i-1})^{\delta} + \beta\, ln\, \psi_{i-1}$
BCACD	E W G	$ln\psi_i = \omega + a\, (\varepsilon_{i-1})^{\delta} + \beta\, ln\, \psi_{i-1}$
STV-BCACD - STD-BCACD	E W G	$ln\psi_i = \omega + a\, (\varepsilon_{i-1})^{\delta'} + \beta\, ln\, \psi_{i-1}$ $\delta' = \delta_1 * \big(1 - G(S_i{:}\, g, s)\big) + \delta_2 * G(S_i{:}\, g, s)$

E, W and G stand for Exponential, Weibull and Generalized Gamma distributions. All the models are abbreviated as in presented above. For parsimony reasons all lags have been restricted to 1.

Table 4.2: Functions $g(\varepsilon_i)$ for the innovations, for the long-term forecasts

Models	$\boldsymbol{g(\varepsilon_i)}$
Log-ACD	$aln(\varepsilon_i)$
EXACD	$a\varepsilon_i + \delta\lvert\varepsilon_i - 1\rvert$
BCACD-OTC,	$(a + \zeta * D_{i-1}){\varepsilon_i}^{\delta}$
BCACD,	$(a\varepsilon_i)^{\delta}$
ST-BCACD	$(a\varepsilon_i)^{\{\delta_1(1-G(S_i:g,s))+\delta_2 G(S_i:g,s)\}}$

Table 2 presents the non-linear functions of standardized durations, used in the calculation of 10 steps forecasts.

Appendix 4.B

Tables

Estimation-Maximum Likelihood

Tables 4.3, 4.4, 4.5 and 4.6 summarize the estimation results of all models for ECX I, ECX II, NP I and NP II, respectively. Panel A in each table presents the estimation results, the associated statistics and the hypotheses tests of three existing models: ACD (Eq. (4.4)), Log-ACD (Eq. (4.5)) and EXACD (Eq. (4.7)). Panel B in each table presents similar information for the ACD-OTC family of models, as in Eqs. (4.12) and (4.14). Panel C in each table presents similar information for the ST-BCACD family of models. The first three columns report the estimation results for the BCACD model, as in Eq. (4.8), the next three columns refer to the SEST-BCACD, while the last three columns refer to the STV-BCACD, as in Eqs. (4.15), (4.16), (4.17) and (4.18). E, W and G stand for Exponential, Weibull and Generalized Gamma distributions and γ and λ are the additional parameters. In addition, the first section in each table presents the estimates of the parameters of the models examined, where the values in parentheses are the associated *t-statistics*. The next section presents the Log-Likelihood function value, *L*, and the Bayesian Information Criterion, *BIC*. The next section presents the hypothesis testing for the additional parameters in W and G models, where the values in parentheses are the associated *p-values*. The last section presents the Kolmogorov-Smirnov statistic, *KS-stat*, and the associated *p-values*.

Table 4.3: ECX I-Estimation Results

A. Estimation of Existing Models

	ACD			Log-ACD			EXACD		
Models	E	W	G	E	W	G	E	W	G
Coefficients									
omega	0.0201	0.0287	0.0451	0.0405	0.1348	0.2641	-0.0545	-0.0728	-0.0681
	(8.87)	(8.90)	(11.75)	(2.45)	(13.57)	(36.74)	(-17.43)	(-39.65)	(-34.24)
alpha	0.1238	0.1795	0.2996	0.0367	0.1410	0.2432	0.1432	0.2303	0.3807
	(12.74)	(13.56)	(15.25)	(2.32)	(11.99)	(36.63)	(17.41)	(22.89)	(25.59)
zeta							-0.0932	-0.1723	-0.3240
							(-13.25)	(-14.98)	(-26.92)
beta	0.8586	0.7962	0.6970	0.9548	0.7856	0.6271	0.9751	0.9517	0.9130
	(59.51)	(52.99)	(40.04)	(41.97)	(35.91)	(46.32)	(55.42)	(32.93)	(47.32)
γ		0.6640	0.3039		0.6511	0.1647		0.6632	0.2669
		(26.47)	(37.77)		(24.31)	(11.30)		(26.44)	(36.88)
λ			4.1684			13.8026			5.4276
			(20.89)			(49.46)			(19.84)
L	-29515.01	-21576.69	-20828.688	-30726.29	-21816.49	-20230.44	-29494.59	-21456.77	-20493.14
BIC	1.3862	1.0138	0.9790	1.4431	1.0251	0.9509	1.3855	1.0085	0.9635
H(0)									
zeta=0							175.69	224.31	42816.38
							(0.00)	(0.00)	(0.00)
$\gamma=1$		22497.11	7484.90		21607.79	318460.06		19699.77	856444.51
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
$\lambda=1$			252.10			2104.90			19196.455
			(0.00)			(0.00)			(0.00)
$\gamma=\lambda=1$			244282			3424821.4			9384582
			(0.00)			(0.00)			(0.00)
Fitting									
KS-stat	0.0086	0.0072	0.0044	0.0069	0.0061	0.0049	0.0123	0.0110	0.0058
$cv_{0.01}=0.0074$	(0.00)	(0.01)	(0.31)	(0.02)	(0.06)	(0.19)	(0.00)	(0.00)	(0.07)

B. ACD-OTC Models

	ACD-OTC			BCACD-OTC		
Models	E	W	G	E	W	G
Coefficients						
omega	0.0266	0.0373	0.0615	-0.1711	-0.3463	-1.3651
	(7.71)	(9.90)	(13.86)	(-6.96)	(-10.22)	(-10.20)
alpha	0.1007	0.1412	0.2053	0.1790	0.3702	1.4859
	(9.84)	(12.94)	(17.52)	(6.70)	(9.83)	(10.65)
zeta	0.0509	0.0736	0.1050	0.0595	0.0992	0.2801
	(8.58)	(10.88)	(14.09)	(7.06)	(10.38)	(20.92)
delta				0.5915	0.4501	0.1816
				(21.38)	(21.10)	(12.37)
beta	0.8485	0.7852	0.6897	0.9653	0.9335	0.8384
	(57.32)	(49.53)	(42.73)	(69.90)	(60.31)	(56.35)
γ		0.6670	0.3157		0.6637	0.1965
		(25.93)	(43.46)		(29.04)	(38.64)
λ			3.9865			10.9840
			(23.81)			(19.65)
L	-29261.30	-21433.42	-20668.39	-29275.00	-21291.33	-19931.86
BIC	1.3746	1.0074	0.9717	1.3755	1.0010	0.9374
H(0)						
zeta=0	73.57	118.44	198.40	49.88	107.70	437.48
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
delta=1				218.09	664.90	3105.53
				(0.00)	(0.00)	(0.00)
$\gamma=1$		18982.77	8868.60		21451.08	10704.36
		(0.00)	(0.00)		(0.00)	(0.00)
$\lambda=1$			318.05			550.89
			(0.00)			(0.00)
$\gamma=\lambda=1$			345251			541689
			(0.00)			(0.00)
Fitting						
KS-stat	0.0063	0.0048	0.0042	0.0158	0.0071	0.0051
$cv_{0.01}=0.0074$	(0.05)	(0.21)	(0.36)	(0.00)	(0.05)	(0.22)

C. Smooth Transition BCACD Models

	BCACD			SEST-BCACD			STV-BCACD		
Models	E	W	G	E	W	G	E	W	G
Coefficients									
omega	-0.1783	-0.3741	-1.8384	-0.1290	-0.2535	-1.7009	-0.1347	-0.2810	-2.2934
	(-6.92)	(-9.80)	(-8.54)	(-9.19)	(-6.85)	(-5.41)	(-6.69)	(-8.55)	(-3.62)
alpha	0.2174	0.4496	2.0691	0.1905	0.3543	1.9353	0.1937	0.3755	2.5167
	(6.84)	(10.11)	(9.24)	(8.81)	(7.61)	(6.06)	(6.92)	(9.69)	(3.38)
delta	0.5434	0.4081	0.1332						
	(21.21)	(19.86)	(9.72)						
delta 1				0.4707	0.3928	0.1286	1.1861	0.6833	0.1466
				(16.28)	(19.54)	(6.92)	(19.77)	(9.98)	(6.08)
delta2				1.2394	0.8003	0.1463	0.4935	0.4082	0.1031
				(11.86)	(9.07)	(5.49)	(5.04)	(21.33)	(15.76)
beta	0.9740	0.9466	0.8837	0.9877	0.9647	0.8842	0.9828	0.9570	0.8829
	(92.57)	(55.00)	(30.17)	(51.50)	(43.21)	(27.80)	(58.14)	(47.81)	(30.30)
g				4.5118	6.1052	5.9850	4.2546	7.3534	6.9540
				(2.91)	(2.44)	(2.81)	(0.98)	(2.15)	(2.78)
s				1.0625	1.0150	1.0045	0.7783	0.8838	0.9065
				(11.55)	(7.19)	(8.69)	(7.33)	(5.97)	(6.89)
γ		0.6627	0.1924		0.6650	0.1932		0.6640	0.1902
		(23.06)	(18.61)		(29.37)	(20.08)		(25.23)	(16.89)
λ			10.2259			10.1474			10.4548
			(9.83)			(10.24)			(8.98)
L	-29465.09	-21401.73	-20181.12	-29226.83	-21347.74	-20180.85	-29329.07	-21374.10	-20179.66
BIC	1.3841	1.0059	0.9488	1.3737	1.0041	0.9491	1.3785	1.0053	0.9490
H(0)									
delta=1	317.71	829.60	2063.44						
	(0.00)	(0.00)	(0.00)						
γ=1		16591.97	7364.36		15782.26	7032.85		20825.08	5170.96
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
λ=1			90.42			85.21			65.91
			(0.00)			(0.00)			(0.00)
γ=λ=1			1201960			1232191			1120516
			(0.00)			(0.00)			(0.00)
Fitting									
KS-stat	0.0101	0.0059	0.0054	0.0079	0.0057	0.0054	0.0056	0.0049	0.0043
$cv_{0.01}$=0.0074	(0.00)	(0.06)	(0.12)	(0.00)	(0.04)	(0.06)	(0.05)	(0.11)	(0.21)

Table 4.4: ECX II-Estimation Results

A. Existing Models

	ACD			Log-ACD			EXACD		
Models	E	W	G	E	W	G	E	W	G
Coefficients									
omega	0.0486	0.0659	0.1107	0.1518	0.1781	0.2483	-0.0654	-0.0764	-0.0551
	(10.04)	(14.41)	(21.15)	(28.11)	(42.49)	(53.81)	(-22.03)	(-36.70)	(-15.56)
alpha	0.1755	0.2381	0.3793	0.1483	0.1889	0.2414	0.2302	0.3160	0.4401
	(18.75)	(24.92)	(33.62)	(25.50)	(39.19)	(56.64)	(20.74)	(62.12)	(35.03)
zeta							-0.1690	-0.2578	-0.3925
							(-15.77)	(-46.52)	(-35.16)
beta	0.7892	0.7114	0.5853	0.7234	0.6618	0.5812	0.9095	0.8731	0.8195
	(62.22)	(58.06)	(49.26)	(49.41)	(58.76)	(58.21)	(63.30)	(56.11)	(40.42)
γ		0.6577	0.1787		0.6581	0.1990		0.6606	0.1990
		(46.16)	(44.46)		(45.04)	(43.13)		(36.42)	(44.27)
λ			10.4856			12.4657			9.7654
			(29.49)			(19.47)			(24.47)
L	-78630.22	-60400.79	-56690.16	-78314.85	-59709.62	-55187.00	-77763.556	-59740.54	-55680.46
BIC	1.7235	1.3242	1.2430	1.7166	1.3090	1.2100	1.7046	1.3098	1.2210
H(0)									
zeta=0							248.77	2163.67	2516.44
							(0.00)	(0.00)	(0.00)
$\gamma=1$		49177.57	3807033		55876.60	3941314		39415.72	3956501
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
$\lambda=1$			1218.09			1759.71			1409.43
			(0.00)			(0.00)			(0.00)
$\gamma=\lambda=1$			3807172			3959833			4835595
			(0.00)			(0.00)			(0.00)
Fitting									
KS-stat	0.0041	0.0029	0.0027	0.0055	0.0054	0.0035	0.0142	0.0130	0.0103
$cv_{0.01}=0.0048$	(0.05)	(0.20)	(0.24)	(0.00)	(0.00)	(0.12)	(0.00)	(0.00)	(0.00)

B. ACD-OTC Models

	ACD-OTC			BCACD-OTC		
Models	E	W	G	E	W	G
Coefficients						
omega	0.0534	0.0775	0.1378	-0.3649	-0.6964	-2.0810
	(10.14)	(15.29)	(22.98)	(-12.77)	(-13.31)	(-9.47)
alpha	0.1672	0.2167	0.2932	0.4420	0.7960	2.2348
	(19.41)	(26.46)	(36.36)	(13.22)	(14.03)	(10.03)
zeta	0.0463	0.0881	0.1442	0.0437	0.1167	0.3203
	(7.12)	(13.57)	(19.26)	(5.44)	(15.03)	(31.07)
delta				0.4068	0.2813	0.1190
				(20.83)	(16.64)	(9.85)
beta	0.7814	0.6952	0.5626	0.9012	0.8518	0.7634
	(59.02)	(55.72)	(46.51)	(60.61)	(57.43)	(42.55)
γ		0.6584	0.1986		0.6609	0.1999
		(43.26)	(59.22)		(49.90)	(51.76)
λ			10.1968			9.9298
			(26.50)			(34.48)
L	-78508.07	-60237.19	-56413.56	-77377.90	-59381.01	-54664.66
BIC	1.7210	1.3207	1.2370	1.6963	1.3021	1.1988
H(0)						
zeta=0	50.72055	184.09	370.97	29.58	225.85	965.10
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
delta=1				923.02	1806.61	5317.06
				(0.00)	(0.00)	(0.00)
$\gamma=1$		43792.66	4221307		63209.93	5057331
		(0.00)	(0.00)		(0.00)	(0.00)
$\lambda=1$			1307.64			1010.50
			(0.00)			(0.00)
$\gamma=\lambda=1$			4226338			5057332
			(0.00)			(0.00)
Fitting						
KS-stat	0.0053	0.0032	0.0028	0.0051	0.0464	0.0042
$cv_{0.01}=0.0048$	(0.00)	(0.15)	(0.19)	(0.00)	(0.04)	(0.05)

C. Smooth Transition BCACD Models

	BCACD			SEST-BCACD			STV-BCACD		
Models	E	W	G	E	W	G	E	W	G
Coefficients									
omega	-0.3778	-0.7564	-3.2812	-0.3091	-0.6258	-2.3955	-0.3114	-0.6053	-2.0981
	(-11.06)	(-12.30)	(-5.26)	(-12.99)	(-11.88)	(-7.27)	(-12.88)	(-12.37)	(-10.96)
alpha	0.4679	0.8889	3.5173	0.3941	0.7580	2.6356	0.3983	0.7385	2.3406
	(11.56)	(13.30)	(5.62)	(13.50)	(13.06)	(7.93)	(13.05)	(13.59)	(12.13)
delta	0.3896	0.2551	0.1174						
	(17.41)	(15.16)	(5.21)						
delta 1				0.3786	0.2523	0.0989	0.5080	0.3381	0.1207
				(20.60)	(15.02)	(0.73)	(17.39)	(16.45)	(10.40)
delta2				0.5107	0.3233	0.2936	0.3844	0.2560	0.0834
				(15.99)	(12.89)	(2.69)	(16.76)	(13.78)	(10.51)
beta	0.9080	0.8718	0.8258	0.9159	0.8771	0.8266	0.9145	0.8771	0.8284
	(67.30)	(54.99)	(53.13)	(63.64)	(51.13)	(63.15)	(71.75)	(40.21)	(60.70)
g				1.5066	2.0784	0.0146	3.3074	2.9474	2.6592
				(1.34)	(1.42)	(1.17)	(1.86)	(2.11)	(1.02)
s				2.1682	2.3865	2.1778	0.4051	0.4517	0.5377
				(11.84)	(23.56)	(12.61)	(13.94)	(19.47)	(8.35)
γ		0.6611	0.1991		0.6614	0.1991		0.6614	0.1992
		(39.46)	(54.47)		(37.20)	(46.36)		(37.25)	(44.50)
λ			11.8453			10.4985			10.3584
			(23.87)			(29.77)			(28.48)
L	-77441.93	-59500.57	-55165.27	-77362.83	-59480.26	-55155.45	-77360.49	-59472.88	-55144.16
BIC	1.6976	1.3045	1.2097	1.6962	1.3045	1.2098	1.6962	1.3043	1.2096
H(0)									
delta=1	743.65	1958.41	4209.12						
	(0.00)	(0.00)	(0.00)						
γ=1		37836.79	5720736		41355.12	3827346		28062.47	3052941
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
λ=1			1044.09			1171.22			1108.49
			(0.00)			(0.00)			(0.00)
γ=λ=1			5734443			3827370			3053508
			(0.00)			(0.00)			(0.00)
Fitting									
KS-stat	0.0050	0.0041	0.0037	0.0060	0.0043	0.0034	0.0072	0.0041	0.0032
$cv_{0.01}$=0.0048	(0.01)	(0.05)	(0.12)	(0.00)	(0.05)	(0.14)	(0.00)	(0.05)	(0.15)

Appendix 4.B

Table 4.5: NP I-Estimation Results

A. Existing Models

	ACD			Log-ACD			EXACD		
Models	E	W	G	E	W	G	E	W	G
Coefficients									
omega	0.0625	0.0674	0.0747	0.1262	0.1660	0.2277	-0.0863	-0.0982	-0.0951
	(3.47)	(4.64)	(12.08)	(5.78)	(8.20)	(10.57)	(-4.87)	(-19.51)	(-5.44)
alpha	0.2116	0.2580	0.3138	0.1248	0.1682	0.2018	0.2516	0.3078	0.3569
	(5.40)	(6.51)	(10.75)	(4.99)	(7.50)	(11.05)	(8.09)	(19.23)	(10.41)
zeta							-0.1807	-0.2318	-0.2759
							(-7.42)	(-14.48)	(-8.32)
beta	0.7334	0.6880	0.6533	0.7378	0.6679	0.6244	0.9074	0.8938	0.8830
	(14.31)	(15.18)	(33.70)	(10.61)	(12.40)	(13.29)	(31.55)	(39.31)	(39.44)
γ		0.6253	0.3798		0.6215	0.3236		0.6293	0.3888
		(17.23)	(15.75)		(7.53)	(24.19)		(27.85)	(17.10)
λ			2.3997			3.2261			2.3381
			(19.31)			(13.07)			(9.91)
L	-2853.21	-1940.95	-1911.35	-2901.99	-1941.74	-1893.54	-2810.67	-1920.31	-1891.13
BIC	1.5066	1.0291	1.0158	1.5323	1.0296	1.0064	1.4864	1.0205	1.0073
H(0)									
zeta=0							55.03	209.76	69.19
							(0.00)	(0.00)	(0.00)
$\gamma=1$		3393.76	56636.23		2841.40	2558.39		3323.59	722.69
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
$\lambda=1$			1634.20			81.38			32.17
			(0.00)			(0.00)			(0.00)
$\gamma=\lambda=1$			83218.83			20773.67			16032.65
			(0.00)			(0.00)			(0.00)
Fitting									
KS-stat	0.0276	0.0232	0.0141	0.0222	0.0197	0.0159	0.0398	0.0354	0.0186
$cv_{0.01}=0.0239$	(0.00)	(0.01)	(0.31)	(0.02)	(0.05)	(0.19)	(0.00)	(0.00)	(0.08)

B. ACD-OTC Models

	ACD-OTC			BCACD-OTC		
Models	E	W	G	E	W	G
Coefficients						
omega	0.0689	0.0782	0.0928	-0.4323	-0.7010	-1.2033
	(3.44)	(4.56)	(5.12)	(-3.54)	(-11.51)	(-3.94)
alpha	0.1764	0.2123	0.2453	0.4606	0.7308	1.2589
	(5.45)	(7.11)	(6.78)	(3.47)	(9.55)	(3.99)
zeta	0.0645	0.1023	0.1273	0.1119	0.1831	0.2897
	(2.13)	(3.56)	(3.91)	(2.52)	(5.80)	(6.37)
delta				0.3823	0.2855	0.1869
				(5.00)	(11.95)	(3.92)
beta	0.7264	0.6690	0.6275	0.8705	0.8322	0.7914
	(13.60)	(13.61)	(13.78)	(25.41)	(40.63)	(30.76)
γ		0.6257	0.3783		0.6286	0.3257
		(15.92)	(16.86)		(11.98)	(12.25)
λ			2.4350			3.2336
			(9.79)			(6.69)
L	-2846.32	-1935.87	-1904.16	-2801.31	-1907.26	-1863.19
BIC	1.5052	1.0286	1.0141	1.4837	1.0158	0.9948
H(0)						
zeta=0	18.87	12.65	15.31	16.33	19.1594	40.57
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
delta=1				65.17	94.47	290.47
				(0.00)	(0.00)	(0.00)
$\gamma=1$		4014.62	767.68		3360.72	643.73
		(0.00)	(0.00)		(0.00)	(0.00)
$\lambda=1$			33.30			21.34
			(0.00)			(0.00)
$\gamma=\lambda=1$			18860.22			25766.64
			(0.00)			(0.00)
Fitting						
KS-stat	0.0203	0.0154	0.0136	0.0338	0.0207	0.0144
$cv_{0.01}=0.0239$	(0.06)	(0.22)	(0.36)	(0.00)	(0.06)	(0.29)

C. Smooth Transition BCACD Models

	BCACD			SEST-BCACD			STV-BCACD		
Models	E	W	G	E	W	G	E	W	G
Coefficients									
omega	-0.3959	-0.6100	-0.9445	-0.3792	-0.5840	-0.9782	-0.3395	-0.5620	-0.9236
	(-3.68)	(-3.95)	(-3.05)	(-4.09)	(-3.88)	(-18.47)	(3.82)	(-3.06)	(-4.68)
alpha	0.4775	0.7270	1.1293	0.4635	0.7030	1.1608	0.4433	0.6904	1.1111
	(3.74)	(4.11)	(3.36)	(4.14)	(4.08)	(14.47)	(4.31)	(3.60)	(5.15)
delta	0.4145	0.3272	0.2384						
	(5.76)	(4.77)	(3.47)						
delta 1				0.3535	0.2986	0.1399	0.6013	0.3801	0.3645
				(4.81)	(4.25)	(2.23)	(6.11)	(5.08)	(2.88)
delta2				0.4571	0.3528	0.3281	0.3973	0.3248	0.2266
				(5.68)	(4.62)	(6.07)	(3.15)	(2.66)	(2.09)
beta	0.8943	0.8696	0.8482	0.8969	0.8715	0.8474	0.8976	0.8720	0.8489
	(30.43)	(31.70)	(32.39)	(30.54)	(30.29)	(41.51)	(32.23)	(31.95)	(32.98)
g				10.5006	9.4302	0.0314	3.8299	5.3254	1.1698
				(1.70)	(1.64)	(0.28)	(2.19)	(1.79)	(0.81)
s				1.5875	2.3260	2.9797	0.7198	0.7577	0.8038
				(8.92)	(25.86)	(0.96)	(2.39)	(3.86)	(8.35)
γ		0.6280	0.3534		0.6282	0.3527		0.6283	0.3543
		(14.27)	(12.59)		(9.50)	(13.74)		(18.44)	(16.24)
λ			2.7681			2.7768			2.7560
			(6.89)			(7.69)			(8.96)
L	-2814.23	-1917.49	-1880.80	-2811.68	-1917.07	-1880.76	-2811.75	-1920.12	-1883.79
BIC	1.4883	1.0190	1.0019	1.4934	1.0253	1.0083	1.4935	1.0269	1.0099
H(0)									
delta=1	66.20	22.74	12.06						
	(0.00)	(0.00)	(0.00)						
$\gamma=1$		3814.98	530.87		3194.98	636.01		2321.16	875.89
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
$\lambda=1$			19.39			24.21			32.57
			(0.00)			(0.00)			(0.00)
$\gamma=\lambda=1$			19322.36			17512.98			13812.93
			(0.00)			(0.00)			(0.00)
Fitting									
KS-stat	0.0326	0.0191	0.0174	0.0282	0.0206	0.0195	0.0200	0.0176	0.0155
$cv_{0.01}=0.0239$	(0.00)	(0.06)	(0.12)	(0.00)	(0.04)	(0.06)	(0.05)	(0.11)	(0.21)

Table 4.6: NP II-Estimation Results

A. Existing Models

	ACD			Log-ACD			EXACD		
Models	E	W	G	E	W	G	E	W	G
Coefficients									
omega	0.1123	0.1329	0.1379	0.1168	0.1402	0.1565	-0.0494	-0.0542	-0.0510
	(2.95)	(2.88)	(3.09)	(6.77)	(7.74)	(6.96)	(-4.06)	(-3.57)	(-2.88)
alpha	0.1263	0.1721	0.1829	0.1133	0.1328	0.1407	0.2372	0.2856	0.2965
	(5.43)	(4.55)	(4.51)	(7.30)	(8.86)	(8.50)	(6.22)	(8.90)	(8.09)
zeta							-0.2007	-0.2467	-0.2580
							(-4.71)	(-7.80)	(-6.64)
beta	0.7619	0.7016	0.6897	0.5931	0.5864	0.5811	0.8206	0.8057	0.8063
	(14.00)	(9.08)	(9.25)	(7.72)	(9.34)	(8.82)	(16.47)	(20.21)	(19.56)
γ		0.6338	0.5782		0.6351	0.5297		0.6382	0.5664
		(8.08)	(15.72)		(8.76)	(14.59)		(8.04)	(15.97)
λ			1.1636			1.3542			1.21953951
			(10.21)			(8.91)			(10.12)
L	-3338.22	-2558.04	-2556.90	-3315.56	-2537.43	-2533.29	-3286.11	-2529.05	-2527.11
BIC	1.8583	1.4279	1.4295	1.8457	1.4164	1.4164	1.8317	1.4140	1.4152
H(0)									
zeta=0							22.15	60.78	44.04
							(0.00)	(0.00)	(0.00)
$\gamma=1$		2588.95	131.42		2316.44	167.70		2324.05	149.54
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
$\lambda=1$			2.06			5.44			3.32
			(0.15)			(0.02)			(0.07)
$\gamma=\lambda=1$			2907.79			4164.64			1523.05
			(0.00)			(0.00)			(0.00)
Fitting									
KS-stat	0.0256	0.0181	0.0170	0.0277	0.0271	0.0175	0.0717	0.0657	0.0521
$cv_{0.01}=0.0242$	(0.01)	(0.20)	(0.19)	(0.00)	(0.00)	(0.12)	(0.00)	(0.00)	(0.00)

B. ACD-OTC Models

	ACD-OTC			BCACD-OTC		
Models	E	W	G	E	W	G
Coefficients						
omega	0.1154	0.1371	0.1432	-0.4843	-0.7557	-0.9202
	(2.50)	(3.88)	(3.51)	(-2.35)	(-3.23)	(-2.88)
alpha	0.1117	0.1399	0.1474	0.5441	0.8129	0.9827
	(3.71)	(3.98)	(3.63)	(2.49)	(3.40)	(2.98)
zeta	0.0250	0.0589	0.0689	0.0653	0.1348	0.1686
	(5.98)	(12.04)	(2.35)	(2.45)	(3.89)	(4.90)
delta				0.2662	0.1962	0.1678
				(2.92)	(3.40)	(2.83)
beta	0.7605	0.6999	0.6860	0.7696	0.7416	0.7297
	(10.44)	(11.56)	(9.31)	(10.14)	(12.97)	(15.71)
γ		0.6338	0.5698		0.6368	0.5194
		(14.27)	(17.06)		(9.63)	(15.48)
λ			1.1924			1.4064
			(10.59)			(9.90)
L	-3336.98	-2556.13	-2554.63	-3293.78	-2525.78	-2520.82
BIC	1.8599	1.4291	1.4305	1.8382	1.4145	1.4140
H(0)						
zeta=0	10.96	14.15	86.77	12.11	18.37	25.22773
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
delta=1				65.00	194.57	196.24
				(0.00)	(0.00)	(0.00)
$\gamma=1$		3629.37035	165.94		3230.57	205.16
		(0.00)	(0.00)		(0.00)	(0.00)
$\lambda=1$			2.92			8.18
			(0.09)			(0.00)
$\gamma=\lambda=1$			1291.49			2464.97
			(0.00)			(0.00)
Fitting						
KS-stat	0.0214	0.0131	0.0115	0.0208	0.0189	0.0171
$cv_{0.01}=0.0242$	(0.03)	(0.42)	(0.59)	(0.04)	(0.08)	(0.14)

C. Smooth-Transition BCACD Models

	BCACD			SEST-BCACD			STV-BCACD		
Models	E	W	G	E	W	G	E	W	G
Coefficients									
omega	-0.4563	-0.6776	-0.7738	-0.3002	-0.4743	-0.5364	-0.2987	-0.4278	-0.4796
	(-2.43)	(-2.54)	(-3.60)	(-3.11)	(-3.00)	(-3.16)	(-3.35)	(-3.41)	(-3.36)
alpha	0.5457	0.7972	0.9097	0.4012	0.6093	0.4599	0.4129	0.5818	0.6526
	(2.59)	(2.78)	(3.93)	(3.35)	(3.34)	(3.59)	(3.57)	(3.96)	(3.89)
delta	0.2860	0.2262	0.2064						
	(3.16)	(2.87)	(4.95)						
delta 1				0.2656	0.1998	0.1793	0.6069	0.4801	0.4385
				(3.14)	(2.59)	(3.40)	(3.02)	(3.04)	(3.30)
delta2				0.5512	0.3827	0.3494	0.2523	0.1909	0.1710
				(3.26)	(3.29)	(3.36)	(3.18)	(3.26)	(3.56)
beta	0.7890	0.7727	0.7676	0.8404	0.8069	0.7996	0.8262	0.8049	0.7977
	(10.81)	(14.30)	(12.12)	(13.63)	(13.50)	(15.13)	(14.68)	(14.61)	(14.51)
g				8.5006	9.3534	9.9002	2.4758	2.7121	2.8036
				(3.70)	(2.18)	(2.13)	(3.96)	(2.99)	(2.91)
s				2.6766	3.3260	3.9797	0.6051	0.7326	0.6038
				(17.42)	(15.86)	(13.96)	(13.94)	(15.86)	(8.35)
γ		0.6369	0.5431		0.6374	0.5442		0.6379	0.5463
		(11.80)	(15.10)		(77.20)	(14.97)		(8.64)	(15.00)
λ			1.3042			1.3014			1.2950
			(9.11)			(9.37)			(9.60)
L	-3296.76	-2530.45	-2527.20	-3288.75	-2527.74	-2524.54	-3287.51	-2526.11	-2523.00
BIC	1.8376	1.4148	1.4153	1.8399	1.4201	1.4206	1.8393	1.4192	1.4198
H(0)									
delta=1	62.32	96.25	362.15						
	(0.00)	(0.00)	(0.00)						
γ=1		3368.98	161.37		1928.04	157.26		2531.75	155.12
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
λ=1			4.51			4.71			4.79
			(0.03)			(0.03)			(0.03)
γ=λ=1			1876.12			1734.30			3067.37
			(0.00)			(0.00)			(0.00)
Fitting									
KS-stat	0.0253	0.0208	0.0187	0.0402	0.0289	0.0226	0.0488	0.0273	0.0214
$cv_{0.01}=0.0242$	(0.01)	(0.04)	(0.08)	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.03)

Appendix 4.C

Fitting and Forecasting

Table 4.7 in Appendix 4.C presents the ranking of the models according to their *Q-statistics*. The first column presents the models and the associated distributions. The next three columns report the ranking, the *Q-statistic* and the *p-value* in parenthesis for ECX I. The next three sections present similar results for ECX II, NP I and NP II, respectively. Table 4.8 in Appendix 4C reports the ranking of the models according to their in-sample one-step "forecasts". The first column presents the models along with the associated distributions. The first section reports the rank and the actual value of *UNL* in both markets and phases. The next section reports similar results according to *CORR* loss function. The next two columns present the average ranking of each model in each phase. The next two columns present the average ranking of each model in each market, while the last reports the average ranking (Total) of each model. Table 4.9 presents similar results according to the out-of-sample one-step forecasts of the models, while Table 4.10 focuses on the out-of-sample ten-step forecasts.

Finally figures 4.1, 4.2, 4.3 and 4.4 present the Q-Q plots of all models in ECX I, ECX II, NP I and NP II, respectively. Panel A in each figure presents the Q-Q plots of existing models. The first column refers to the basic ACD model, the second refers to the Log-ACD, while the last to the EX-ACD. Panel B in each figure presents the Q-Q plots of new models proposed in this study. The first two columns refer to the ACD-OTC and BCACD-OTC models, while the last column refers to the STM-ACD, T-ACD and T-ACD-OTC models. Panel C in each figure presents the Q-Q plots of the BCACD (first column), the SEST-BCACD (second column) and the STV-BCACD (third column) models. E, W and G refer to the Exponential, Weibull and Generalized Gamma distributions.

Appendix 4.C

Table 4.7: Q-statistics at 15 lags

		Rank	ECX I	Pr	Rank	ECX II	Pr	Rank	NP I	Pr	Rank	NP II	Pr
ACD	E	4	8.16	(0.88)	9	14.88	(0.46)	1	1.46	(1.00)	3	6.33	(0.97)
	W	3	6.85	(0.94)	7	14.63	(0.48)	4	1.81	(1.00)	5	6.35	(0.97)
	G	1	5.51	(0.99)	5	14.01	(0.51)	7	2.11	(1.00)	4	6.33	(0.97)
Log-ACD	E	26	43.00	(0.00)	22	30.88	(0.00)	27	44.88	(0.00)	27	68.07	(0.00)
	W	27	47.20	(0.00)	14	28.49	(0.00)	26	40.96	(0.00)	26	66.10	(0.00)
	G	25	35.40	(0.00)	12	26.70	(0.00)	25	33.96	(0.00)	25	63.12	(0.00)
EXACD	E	22	30.20	(0.00)	17	29.87	(0.00)	19	19.47	(0.00)	18	23.58	(0.00)
	W	16	24.10	(0.00)	16	29.82	(0.00)	17	15.75	(0.00)	16	22.36	(0.00)
	G	14	17.90	(0.00)	13	28.03	(0.00)	15	10.75	(0.00)	13	20.02	(0.00)
BCACD	E	24	30.80	(0.00)	19	30.03	(0.00)	18	17.26	(0.00)	22	26.54	(0.21)
	W	19	25.30	(0.00)	21	30.68	(0.00)	16	13.06	(0.00)	20	25.20	(0.22)
	G	11	12.40	(0.00)	18	29.98	(0.00)	12	8.44	(0.00)	17	23.06	(0.25)
SEST-BCACD	E	23	30.70	(0.00)	25	40.23	(0.00)	22	21.06	(0.00)	21	25.57	(0.22)
	W	17	24.60	(0.00)	27	41.33	(0.00)	21	20.42	(0.00)	10	14.80	(0.39)
	G	13	17.40	(0.00)	26	40.51	(0.00)	14	9.75	(0.00)	15	21.99	(0.25)
STV-BCACD	E	20	27.80	(0.00)	24	31.68	(0.00)	24	33.08	(0.00)	23	27.76	(0.21)
	W	15	21.00	(0.00)	15	29.28	(0.00)	20	20.29	(0.00)	12	17.80	(0.28)
	G	10	11.75	(0.00)	10	15.64	(0.00)	11	8.17	(0.00)	11	16.17	(0.37)
ACD-OTC	E	9	11.01	(0.75)	8	14.72	(0.46)	2	1.71	(1.00)	7	6.51	(0.97)
	W	7	9.62	(0.84)	6	14.29	(0.48)	5	1.85	(1.00)	8	6.52	(0.97)
	G	6	8.76	(0.89)	4	13.79	(0.51)	8	2.17	(1.00)	6	6.51	(0.97)
BCACD-OTC	E	21	29.90	(0.00)	23	31.46	(0.00)	23	24.07	(0.00)	24	29.97	(0.01)
	W	18	24.70	(0.00)	20	30.11	(0.00)	13	9.68	(0.00)	19	24.21	(0.06)
	G	12	14.70	(0.00)	11	19.66	(0.00)	10	5.88	(0.00)	14	21.58	(0.12)
STM-ACD		8	9.98	(0.00)	3	13.40	(0.00)	3	1.72	(1.00)	1	5.57	(0.99)
T-ACD		2	6.81	(0.96)	1	10.04	(0.82)	6	1.96	(1.00)	9	10.76	(0.77)
T-ACD-OTC		5	8.25	(0.91)	2	10.39	(0.79)	9	2.64	(1.00)	2	5.61	(0.99)

Table 4.8: In-Sample Forecasts

		UNL								CORR								RANK				
		Rank	ECX I	Rank	ECX II	Rank	NP I	Rank	NP II	Rank	ECX I	Rank	ECX II	Rank	NP I	Rank	NP II	Phase I	Phase II	ECX	NP	Total
ACD	E	3	6.6977	14	3.8849	15	2.6001	20	1.2568	9	0.3904	21	0.2085	9	0.3550	9	0.1908	8	16	11	10	10
	W	5	6.7994	18	3.9595	16	2.6100	25	1.2630	7	0.4181	11	0.2241	8	0.3568	8	0.1938	8	14	9	12	9
	G	2	6.5263	11	3.8205	11	2.5203	23	1.2624	2	0.4374	2	0.2320	6	0.3594	6	0.1974	3	7	3	9	7
Log-ACD	E	21	8.2731	9	3.6757	9	2.5012	26	1.2689	24	0.2707	10	0.2243	26	0.2808	18	0.1844	21	15	16	20	18
	W	12	7.7151	8	3.6356	4	2.4390	24	1.2629	23	0.2782	9	0.2262	25	0.2810	16	0.1847	17	13	13	17	15
	G	8	6.8834	2	3.4283	3	2.4112	22	1.2592	22	0.2824	8	0.2275	24	0.2815	15	0.1848	12	10	8	15	12
EXACD	E	23	8.3093	26	4.3735	26	2.6684	19	1.2389	27	0.1699	27	0.1004	27	0.2796	21	0.1829	27	27	26	27	27
	W	27	8.5721	27	4.5028	25	2.6674	18	1.2322	26	0.1724	26	0.1059	19	0.2894	19	0.1834	26	26	27	23	26
	G	26	8.5292	25	4.1261	17	2.6217	17	1.2262	25	0.2335	25	0.1513	17	0.2939	13	0.1859	22	24	25	15	23
BCACD	E	22	8.3026	19	4.0329	24	2.6590	16	1.2205	18	0.3056	19	0.2120	22	0.2856	21	0.1829	23	20	22	24	21
	W	20	8.2333	22	4.0557	21	2.6561	15	1.2124	16	0.3178	17	0.2148	16	0.2940	20	0.1833	19	19	20	18	19
	G	14	7.8488	17	3.9336	14	2.5763	13	1.2052	14	0.3240	18	0.2146	14	0.3009	17	0.1846	11	17	15	13	15
SEST-BCACD	E	25	8.5107	24	4.0683	22	2.6570	11	1.2006	19	0.3048	20	0.2088	23	0.2850	23	0.1805	25	21	24	20	25
	W	24	8.4688	23	4.0596	23	2.6585	21	1.2580	13	0.3259	16	0.2176	15	0.2944	27	0.1396	20	25	21	26	22
	G	15	7.9891	16	3.9076	18	2.6527	27	1.2860	11	0.3328	14	0.2188	13	0.3019	14	0.1853	12	18	14	18	17
STV-BCACD	E	17	8.1441	21	4.0400	27	2.7664	10	1.1848	21	0.2931	23	0.2018	21	0.2880	25	0.1604	23	22	23	24	23
	W	16	8.1374	15	3.9039	19	2.6556	14	1.2079	15	0.3230	24	0.2014	20	0.2884	26	0.1529	18	22	18	20	20
	G	13	7.8237	7	3.6338	20	2.6559	12	1.2018	10	0.3387	12	0.2214	18	0.2930	11	0.1880	15	7	10	14	11
ACD-OTC	E	7	6.8600	12	3.8373	7	2.4833	9	1.1488	8	0.4044	6	0.2280	4	0.3641	5	0.2010	6	6	7	6	6
	W	9	6.9920	5	3.5861	8	2.4947	8	1.1458	6	0.4214	5	0.2289	3	0.3665	4	0.2019	6	4	5	4	4
	G	10	7.0723	6	3.5997	5	2.4475	7	1.1366	4	0.4306	3	0.2319	2	0.3681	2	0.2037	3	3	4	3	3
BCACD-OTC	E	18	8.2077	4	3.5405	13	2.5419	6	1.1225	20	0.3043	22	0.2045	12	0.3046	24	0.1763	16	12	16	11	14
	W	19	8.2135	20	4.0387	12	2.5282	5	1.0935	17	0.3144	15	0.2180	11	0.3093	10	0.1881	14	11	19	8	13
	G	11	7.5186	13	3.8710	10	2.5129	4	1.0743	12	0.3296	13	0.2198	10	0.3096	12	0.1875	10	7	12	7	8
STM-ACD		6	6.8152	10	3.8159	6	2.4539	3	1.0621	5	0.4269	7	0.2277	7	0.3582	7	0.1943	5	5	6	4	5
T-ACD		4	6.7113	1	3.4259	2	2.3379	1	1.0235	3	0.4341	4	0.2316	5	0.3635	3	0.2033	2	2	2	2	2
T-ACD-OTC		1	4.7420	3	3.5100	1	2.2984	2	1.0400	1	0.4545	1	0.2425	1	0.3893	1	0.2176	1	1	1	1	1

Table 4.9: Out-of-Sample Short-Term Forecasts

		UNL								CORR								RANK				
		Rank	ECX I	Rank	ECX II	Rank	NP I	Rank	NP II	Rank	ECX I	Rank	ECX II	Rank	NP I	Rank	NP II	Phase I	Phase II	ECX	NP	Total
ACD	E	20	4.7389	9	4.8883	26	2.6268	15	2.3921	19	0.1501	9	0.2471	20	0.2165	27	0.0448	24	15	13	25	22
	W	21	4.7935	10	4.9699	27	2.6525	12	2.3370	17	0.1561	6	0.2538	18	0.2176	26	0.0463	23	12	10	23	20
	G	19	4.5967	6	4.7732	25	2.5760	11	2.3107	15	0.1604	4	0.2574	17	0.2181	25	0.0469	22	9	7	22	15
Log-ACD	E	5	3.8659	8	4.8167	8	2.0341	20	2.5545	13	0.1620	22	0.2113	10	0.2330	5	0.0940	8	14	8	10	7
	W	4	3.7441	7	4.7792	3	2.0050	18	2.4727	11	0.1637	20	0.2125	9	0.2340	3	0.0941	4	11	6	4	5
	G	3	3.4867	4	4.5142	13	2.1933	17	2.4134	10	0.1640	21	0.2113	7	0.2348	4	0.0941	7	9	5	6	6
EXACD	E	25	18.3581	23	5.4671	24	2.2761	25	2.6501	27	0.0120	27	0.1395	22	0.2061	21	0.0794	26	27	26	27	27
	W	27	25.7743	25	5.5980	22	2.2604	23	2.5937	26	0.0126	26	0.1418	26	0.1967	18	0.0820	27	26	27	26	26
	G	26	24.9345	19	5.3037	18	2.2214	21	2.5675	25	0.0196	25	0.1740	27	0.1859	17	0.0826	25	24	25	23	25
BCACD	E	12	4.1626	17	5.2585	19	2.2252	26	2.6589	24	0.1428	15	0.2222	12	0.2263	20	0.0807	16	22	21	21	22
	W	13	4.1723	21	5.3325	21	2.2536	22	2.5921	23	0.1465	17	0.2178	13	0.2261	15	0.0827	20	20	23	18	22
	G	9	4.0682	16	5.2325	16	2.2200	19	2.5500	9	0.1650	23	0.2088	11	0.2283	14	0.0834	10	19	13	15	13
SEST-BCACD	E	17	4.2367	20	5.3256	23	2.2714	27	2.6795	8	0.1710	16	0.2219	4	0.2424	22	0.0772	11	25	18	20	20
	W	11	4.1581	22	5.3550	20	2.2439	16	2.3941	6	0.1717	18	0.2158	6	0.2360	24	0.0500	9	23	13	16	17
	G	8	4.0680	15	5.2199	15	2.2173	8	2.1778	3	0.1757	24	0.2030	3	0.2425	16	0.0827	5	17	9	7	8
STV-BCACD	E	18	4.3894	18	5.2918	17	2.2210	24	2.6451	7	0.1711	14	0.2234	14	0.2213	19	0.0815	12	20	13	19	18
	W	10	4.1547	14	5.0733	14	2.2002	10	2.2788	12	0.1627	18	0.2158	21	0.2119	23	0.0513	13	18	10	17	15
	G	7	4.0480	5	4.7728	12	2.1070	9	2.1857	4	0.1745	13	0.2254	8	0.2342	13	0.0835	6	5	4	7	4
ACD-OTC	E	23	5.6038	27	5.7783	10	2.0648	14	2.3741	18	0.1513	8	0.2507	19	0.2176	12	0.0875	20	16	24	14	18
	W	24	5.6398	26	5.6552	11	2.0683	13	2.3463	16	0.1569	7	0.2522	16	0.2186	8	0.0906	16	12	22	12	14
	G	22	5.4910	24	5.5278	9	2.0627	4	2.1314	14	0.1605	5	0.2571	15	0.2191	7	0.0919	14	5	20	5	10
BCACD-OTC	E	14	4.2018	12	5.0363	6	2.0212	7	2.1395	22	0.1467	12	0.2313	25	0.2042	11	0.0879	16	8	17	13	12
	W	16	4.2082	13	5.0373	7	2.0213	6	2.1357	21	0.1467	11	0.2318	24	0.2046	10	0.0879	19	5	18	11	11
	G	15	4.2066	11	5.0303	5	2.0212	5	2.1352	20	0.1469	10	0.2318	23	0.2049	9	0.0879	15	4	12	7	9
STM-ACD		6	3.9187	3	4.3941	4	2.0187	3	2.0380	5	0.1725	2	0.2596	5	0.2369	1	0.0999	3	2	3	3	3
T-ACD		2	3.4699	2	4.2631	2	2.0045	2	2.0273	2	0.1976	3	0.2585	2	0.2606	6	0.0922	2	3	2	2	2
T-ACD-OTC		1	3.4695	1	4.2323	1	2.0018	1	2.0174	1	0.1979	1	0.2655	1	0.2698	2	0.0952	1	1	1	1	1

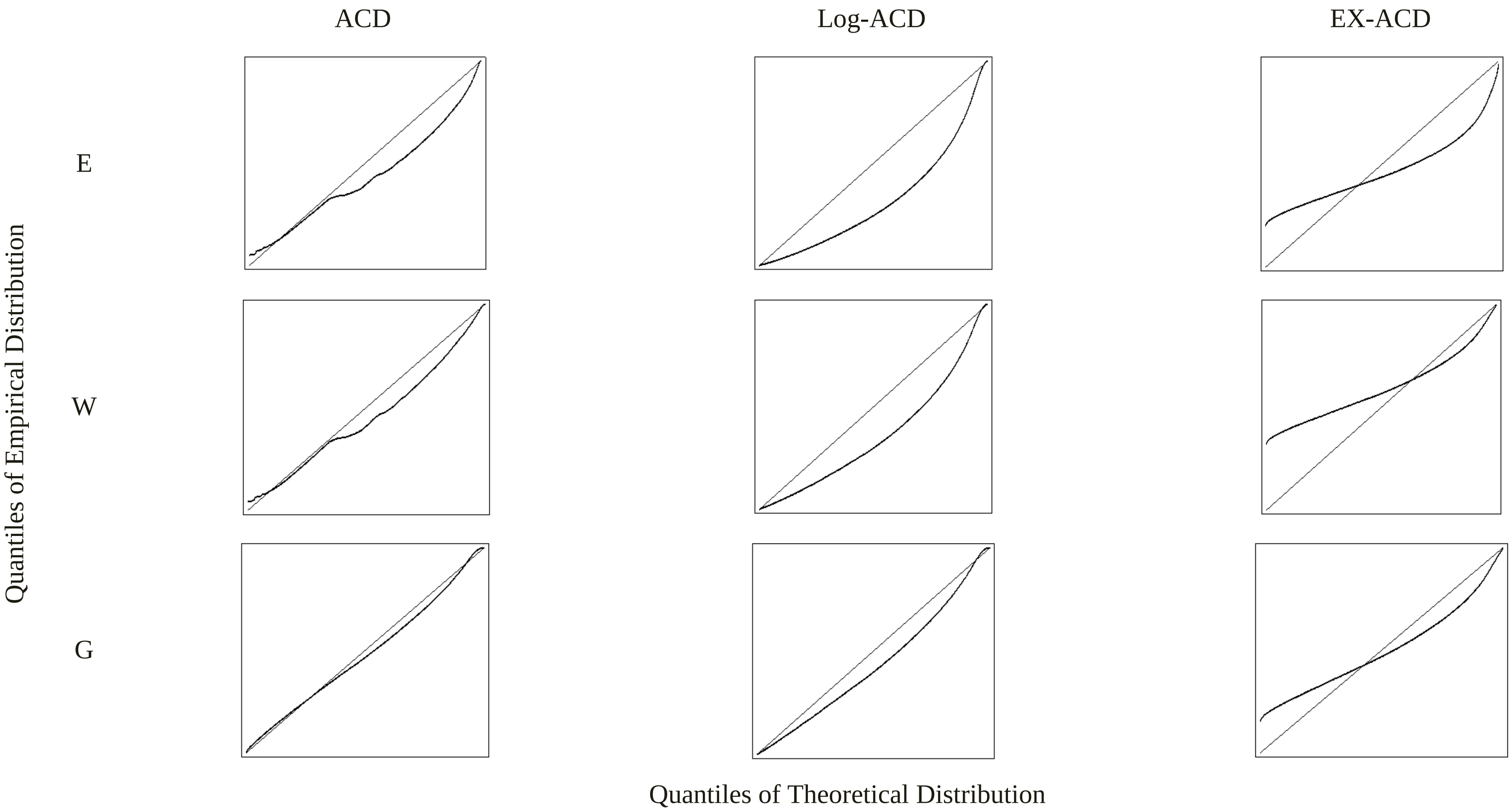
Table 4.10: Out-of-Sample Long-Term Forecasts

		UNL								CORR								RANK				
		Rank	ECX I	Rank	ECX II	Rank	NP I	Rank	NP II	Rank	ECX I	Rank	ECX II	Rank	NP I	Rank	NP II	Phase I	Phase II	ECX	NP	Total
ACD	E	9	4.2328	7	6.3974	9	2.3367	11	2.3641	9	0.0421	3	0.2229	7	0.0438	6	0.0791	5	4	5	6	5
	W	8	4.1868	8	6.5088	8	2.3233	9	2.2582	14	0.0403	6	0.1945	4	0.0480	4	0.0837	5	4	6	4	5
	G	5	3.3839	6	5.7660	7	2.1941	7	2.1805	4	0.0516	5	0.2000	3	0.0488	3	0.0859	3	3	3	3	3
Log-ACD	E	7	3.5080	4	5.3990	5	1.8145	8	2.2104	26	0.0093	20	0.0275	12	0.0373	12	0.0417	9	11	13	10	11
	W	6	3.4609	5	5.4507	4	1.7979	6	2.1308	25	0.0096	19	0.0276	10	0.0398	13	0.0412	8	10	10	6	9
	G	2	2.4434	2	4.9096	3	1.7907	3	2.0182	27	0.0092	17	0.0315	9	0.0429	11	0.0418	7	9	8	5	7
EXACD	E	12	5.3554	26	16.1917	20	5.1700	17	4.4079	24	0.0118	14	0.0468	21	0.0311	10	0.0420	21	17	19	17	18
	W	16	6.2081	22	13.1523	22	5.3994	18	6.9647	23	0.0127	15	0.0403	20	0.0315	14	0.0409	22	18	19	19	20
	G	3	2.5328	16	9.3016	19	5.1621	16	4.1113	22	0.0135	16	0.0340	16	0.0323	9	0.0424	14	14	13	14	14
BCACD	E	24	12.9065	24	13.8989	17	4.8325	22	7.2302	15	0.0402	22	0.0076	27	0.0261	19	-0.0010	26	21	26	23	24
	W	27	14.8943	21	13.0769	24	5.6111	26	8.7122	12	0.0418	25	0.0016	26	0.0261	22	-0.0029	27	24	26	25	27
	G	20	11.1956	20	12.67	27	5.6845	27	8.7157	11	0.0419	26	0.0000	23	0.0297	24	-0.0031	22	26	22	27	25
SEST-BCACD	E	23	12.2849	23	13.2880	16	4.7906	21	7.1995	5	0.0429	18	0.0300	19	0.0316	23	-0.0029	17	20	18	20	19
	W	26	14.5418	25	14.7598	23	5.5456	25	8.4906	8	0.0422	23	0.0029	24	0.0278	26	-0.0041	22	27	25	25	26
	G	19	10.62	18	12.06	26	5.6536	24	8.4810	13	0.0416	27	0.0000	6	0.0441	27	-0.0044	18	25	22	22	22
STV-BCACD	E	22	12.2245	27	16.2424	18	4.8569	19	7.0312	6	0.0427	21	0.0092	25	0.0271	20	-0.0013	19	21	19	21	21
	W	25	14.4614	19	12.5751	25	5.6351	23	7.9753	10	0.0419	24	0.0021	22	0.0298	21	-0.0029	25	21	24	24	23
	G	21	11.87	17	11.37	21	5.1720	20	7.0966	7	0.0425	11	0.0910	5	0.0453	25	-0.0032	11	19	11	18	16
ACD-OTC	E	14	5.8555	11	7.7420	11	2.6845	5	2.0835	18	0.0211	7	0.1198	14	0.0327	8	0.0429	12	8	9	11	9
	W	15	5.9393	12	7.8367	12	2.7301	10	2.3191	20	0.0165	9	0.1042	15	0.0324	15	0.0393	16	12	11	12	12
	G	13	5.7442	10	7.5478	10	2.6698	4	2.0789	16	0.0222	8	0.1099	13	0.0329	7	0.0456	10	6	7	9	8
BCACD-OTC	E	11	4.9309	14	8.1209	14	3.8030	13	3.0570	19	0.0196	13	0.0905	17	0.0322	18	0.0010	15	15	13	15	15
	W	18	10.0790	15	8.4875	15	3.8658	15	3.5977	21	0.0149	12	0.0905	18	0.0321	17	0.0023	20	16	17	16	17
	G	17	8.4774	13	8.0879	13	3.0179	12	3.0422	17	0.0216	10	0.0991	11	0.0382	16	0.0103	13	13	13	12	13
STM-ACD		10	4.5124	9	7.1409	6	1.9705	14	3.3782	3	0.0739	2	0.2297	8	0.0432	5	0.0830	4	7	4	6	4
T-ACD		1	2.2699	3	4.9110	1	1.6376	2	1.8912	1	0.0939	1	0.2356	2	0.0610	2	0.0917	1	2	1	2	1
T-ACD-OTC		4	2.6344	1	4.7169	2	1.7696	1	1.8155	2	0.0921	4	0.2223	1	0.0795	1	0.0958	2	1	2	1	2

Appendix 4.C

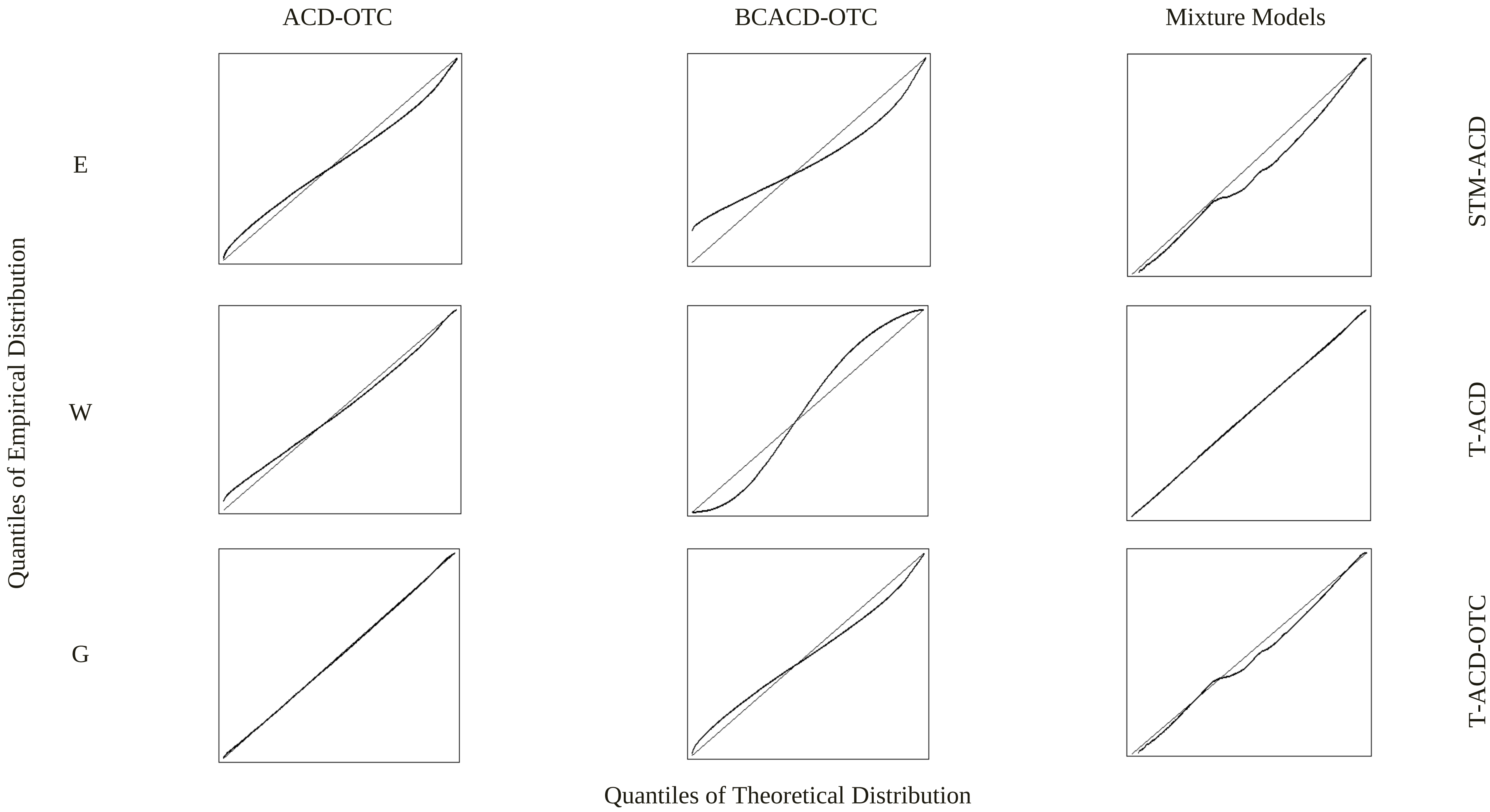
Figure 4.1: Q-Q Plots ECX I

A. Existing Models



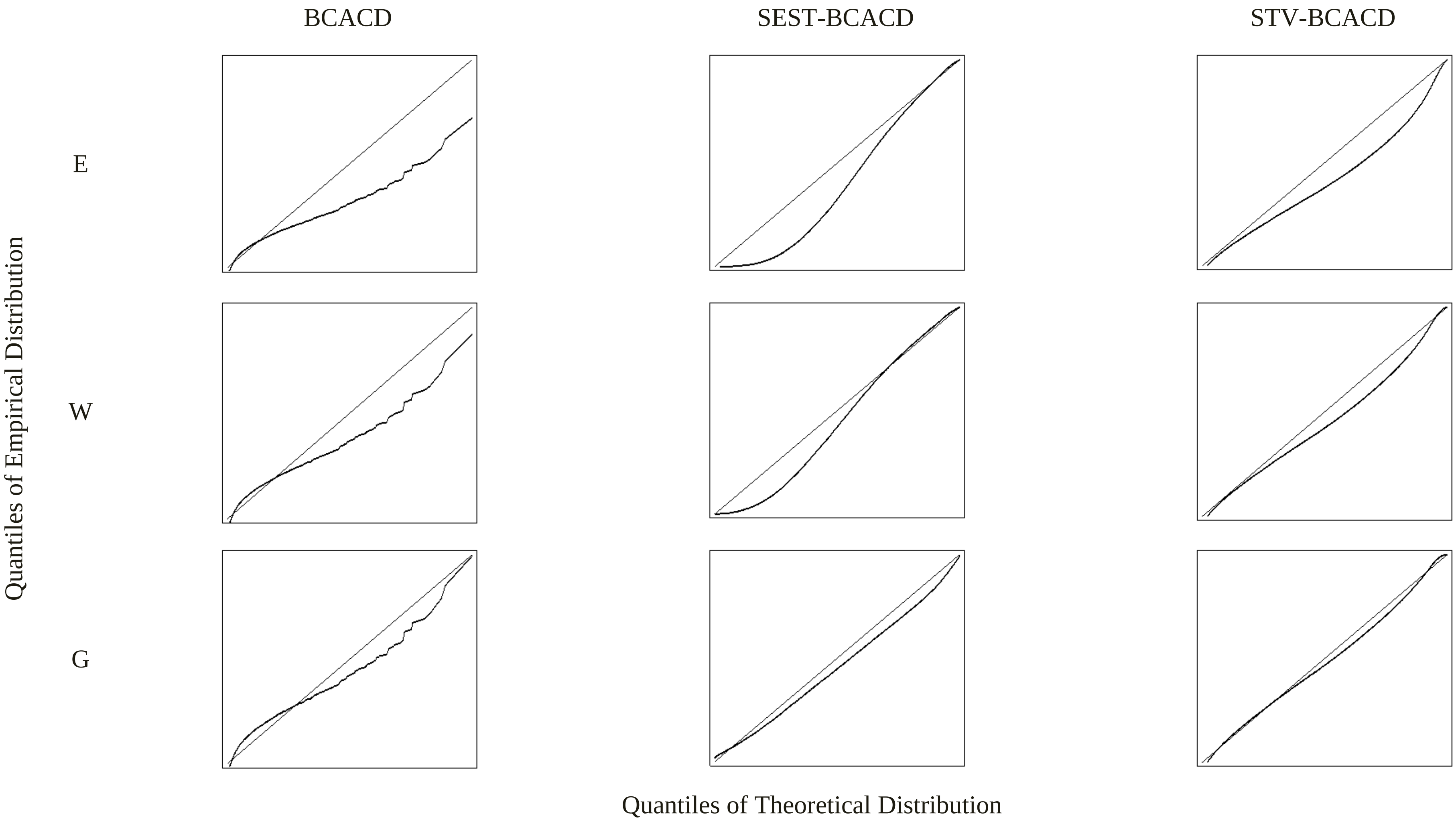
Appendix 4.C

B. ACD-OTC and Mixture of Distributions Models



Appendix 4.C

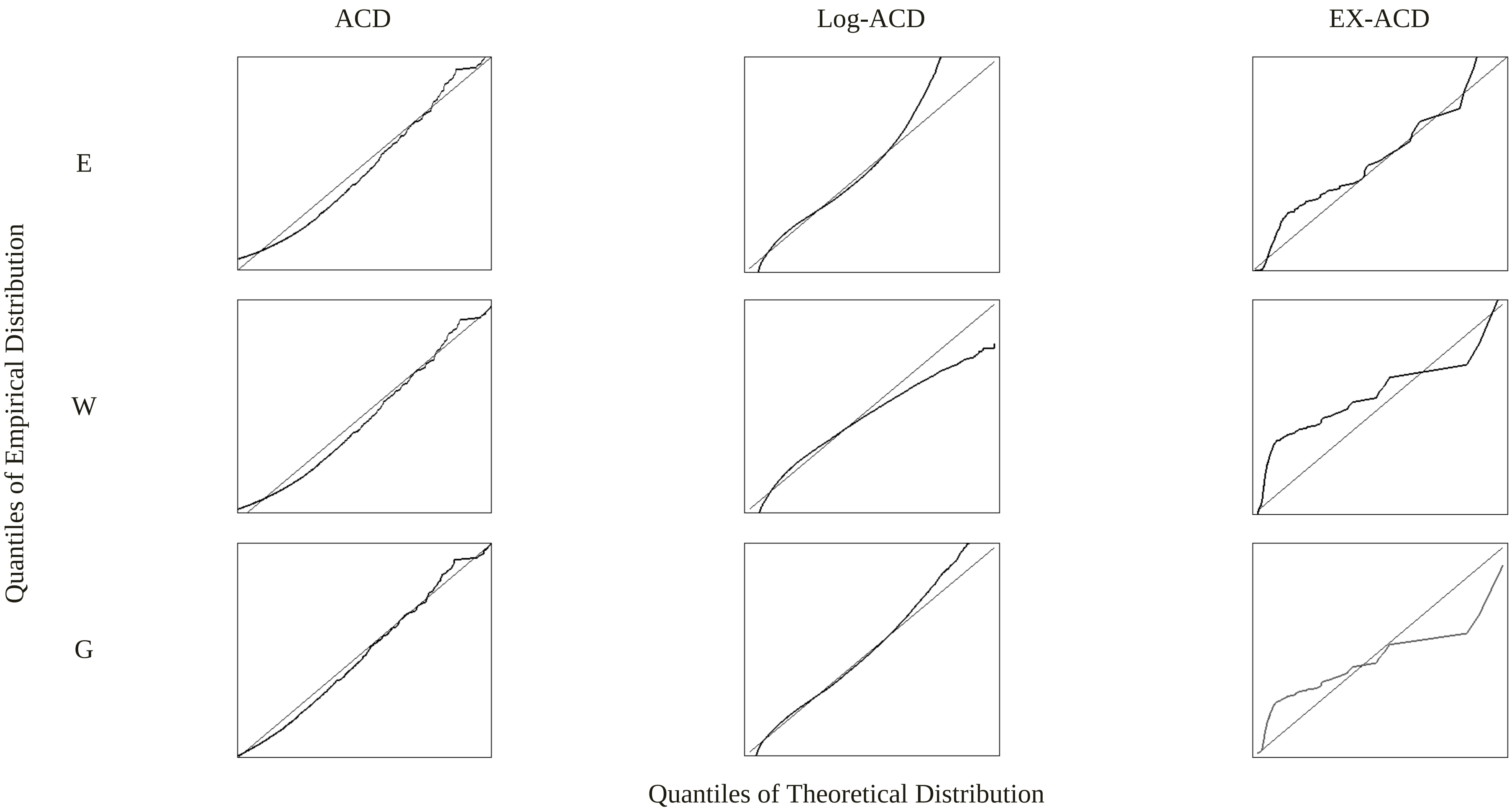
C. ACD-OTC and Mixture of Distributions Models



Appendix 4.C

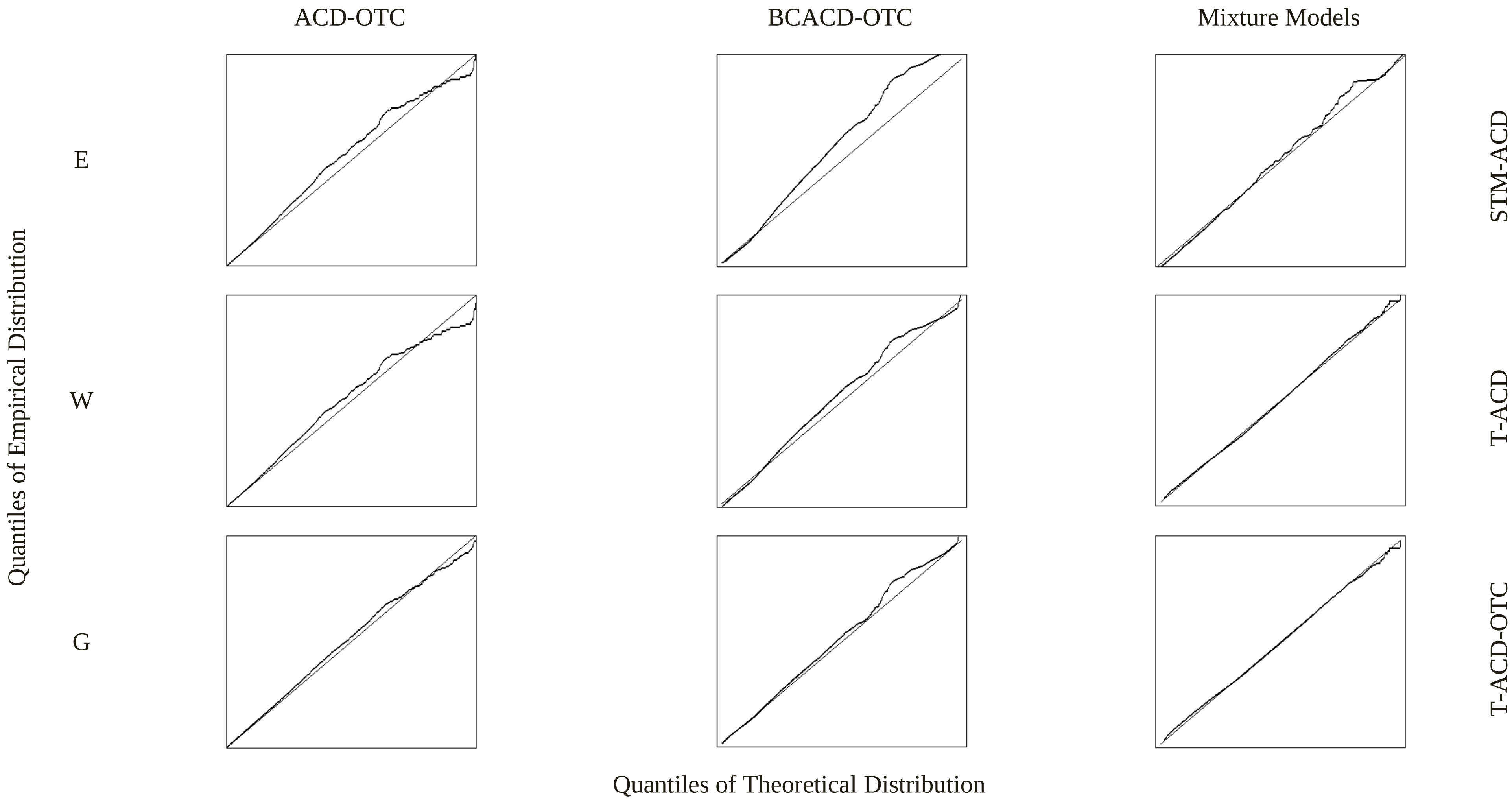
Figure 4.2: Q-Q Plots ECX II

A. Existing Models



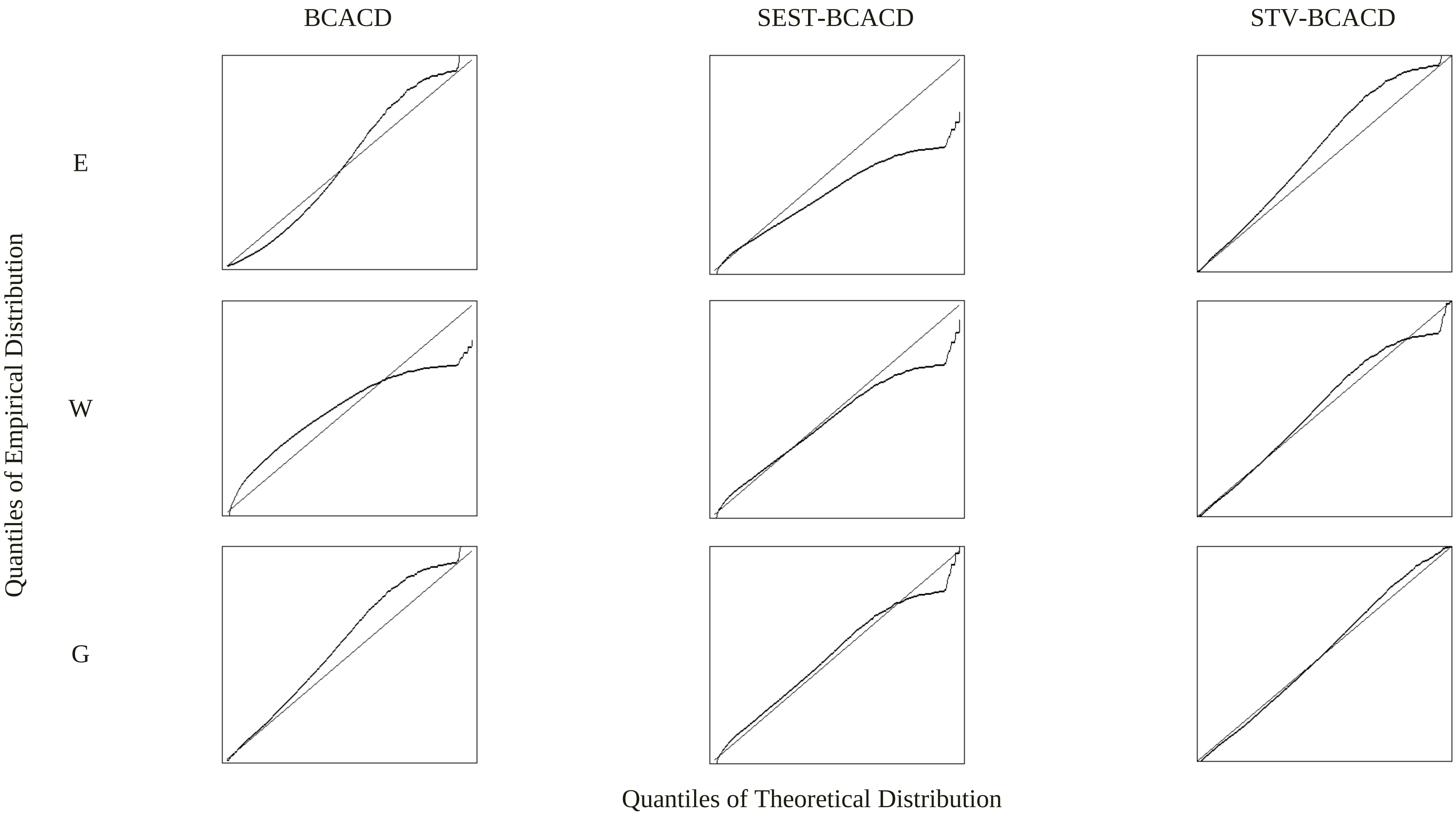
Appendix 4.C

B. ACD-OTC and Mixture of Distributions Models



Appendix 4.C

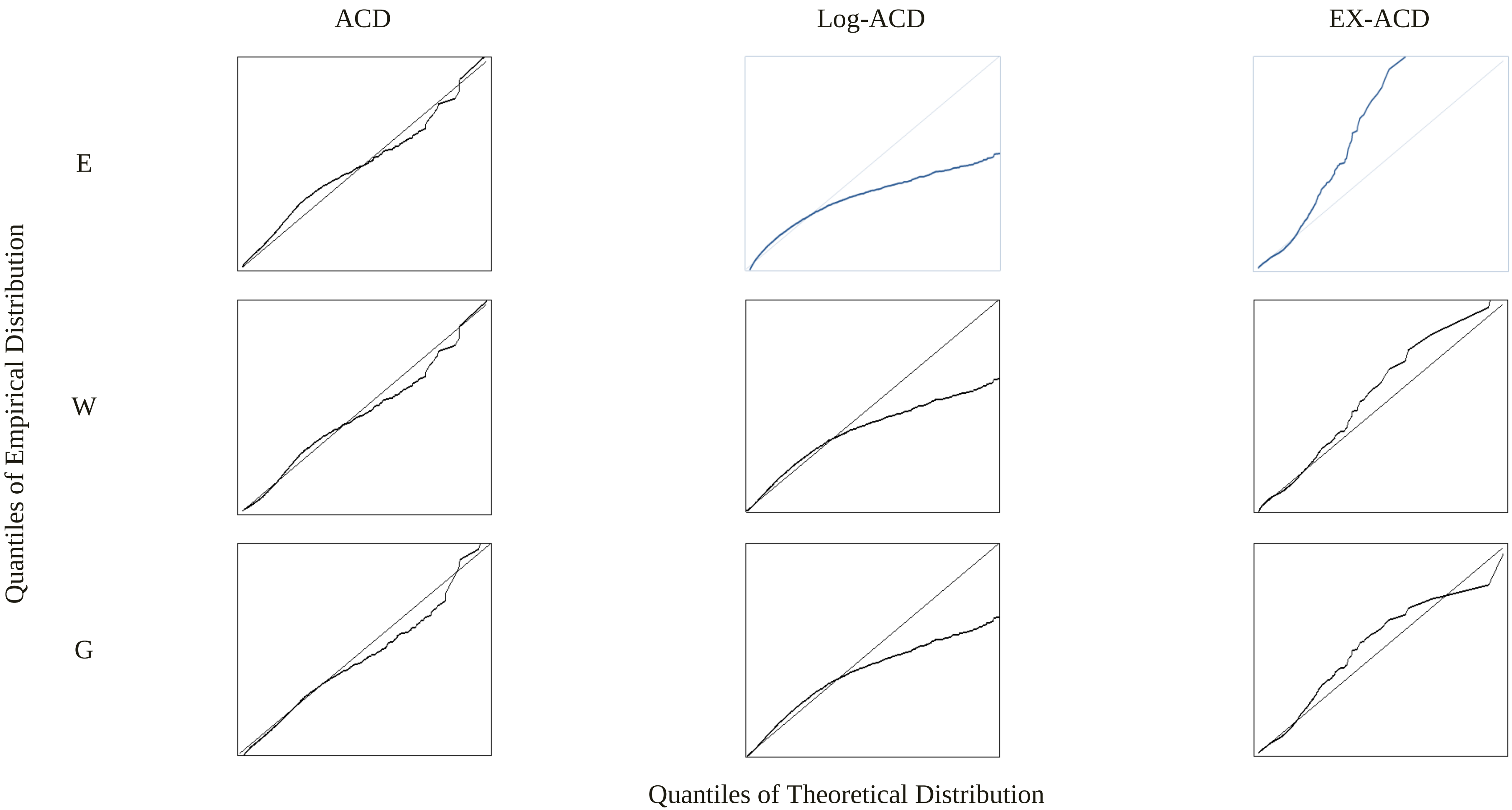
C. ACD-OTC and Mixture of Distributions Models



Appendix 4.C

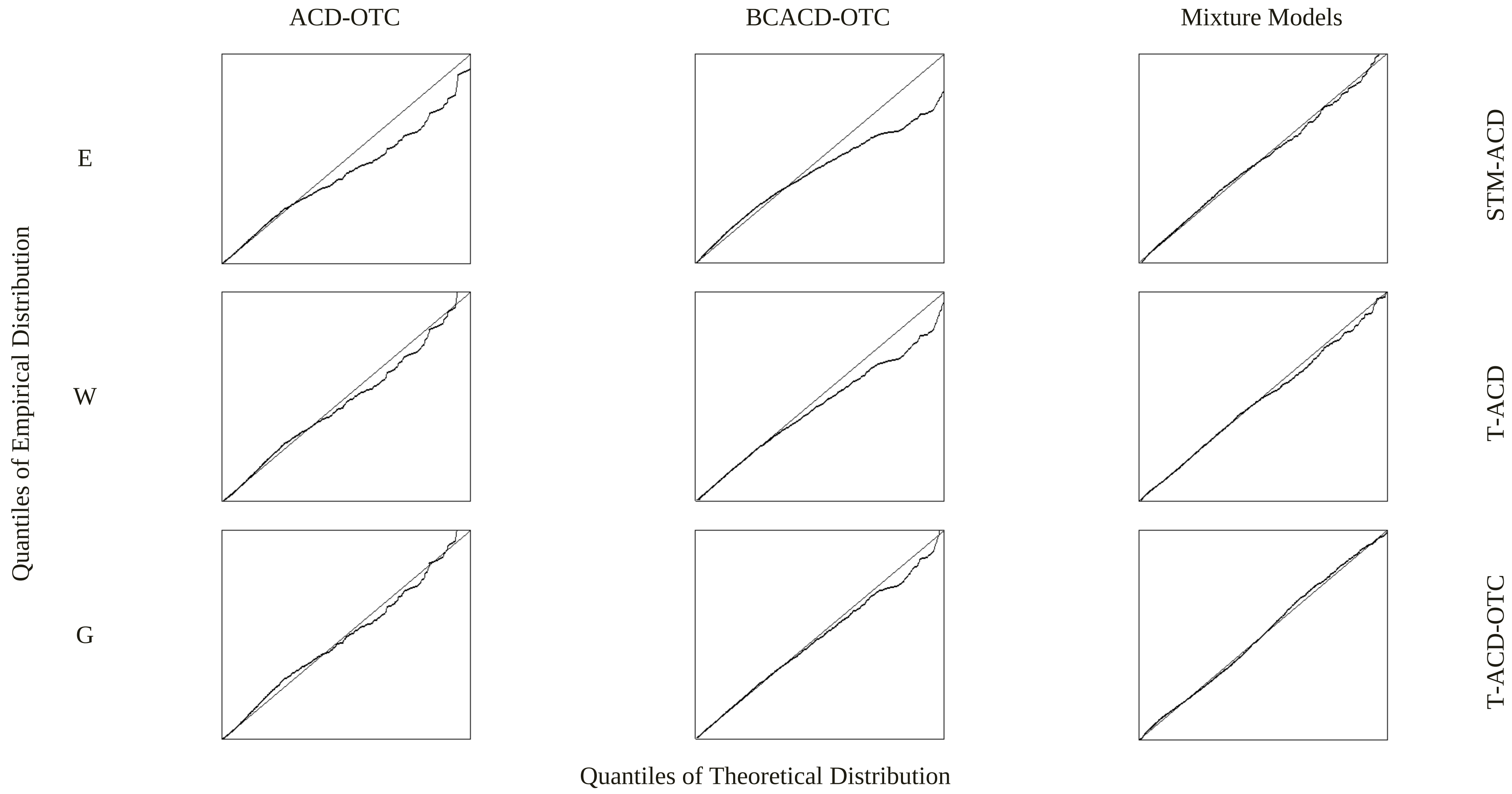
Figure 4.3: Q-Q Plots NP I

A. Existing Models



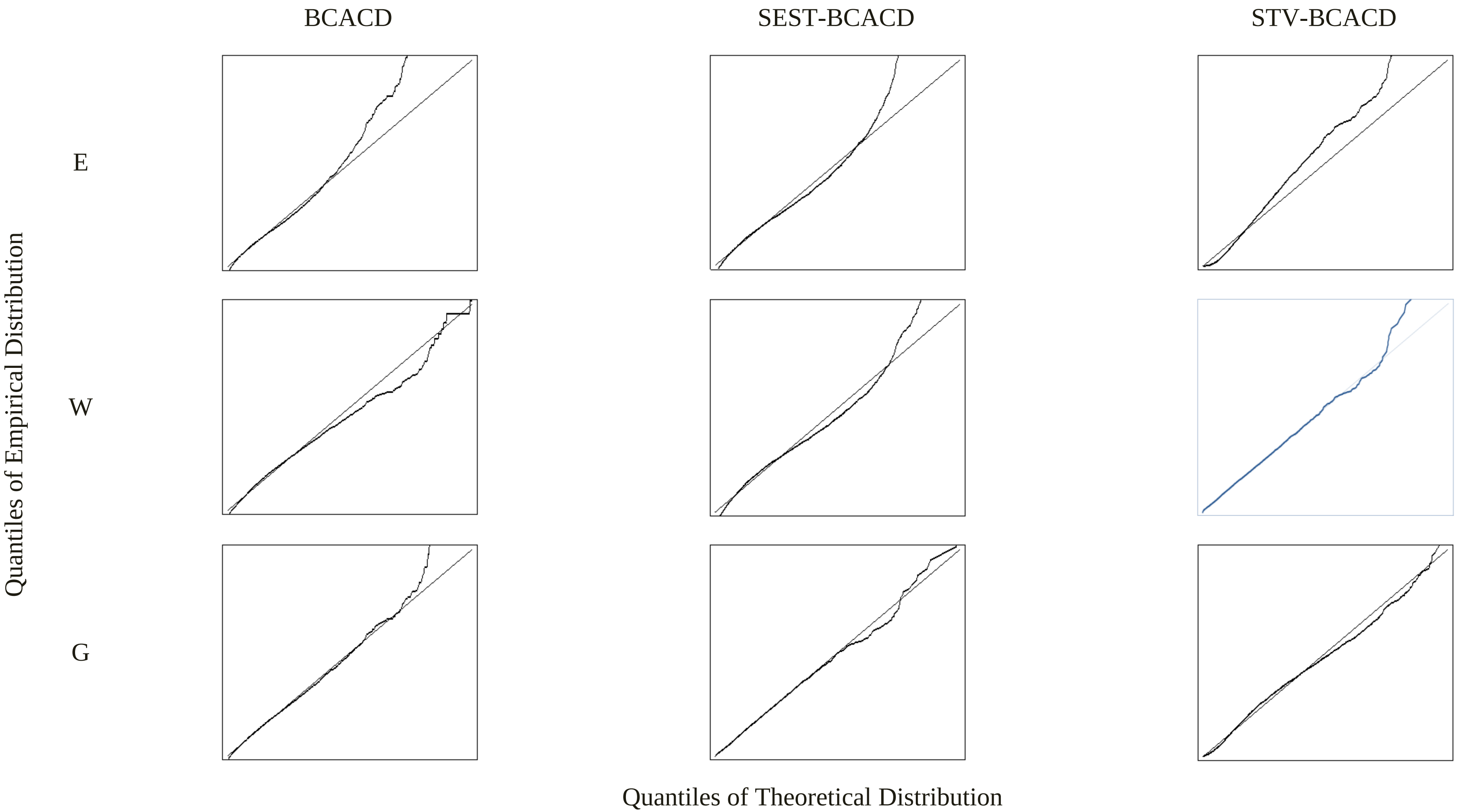
Appendix 4.C

B. ACD-OTC and Mixture of Distributions Models



Appendix 4.C

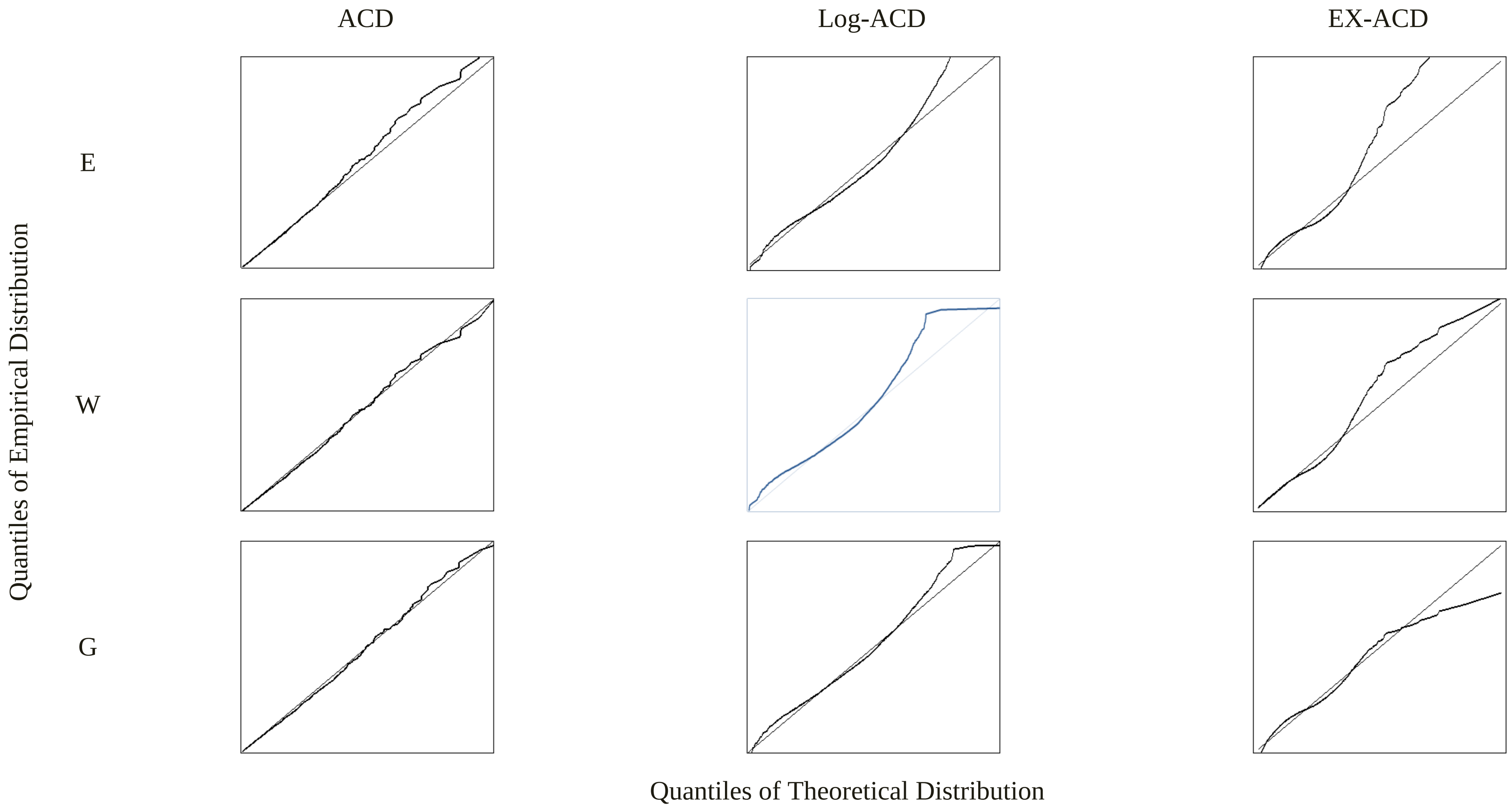
C. ACD-OTC and Mixture of Distributions Models



Appendix 4.C

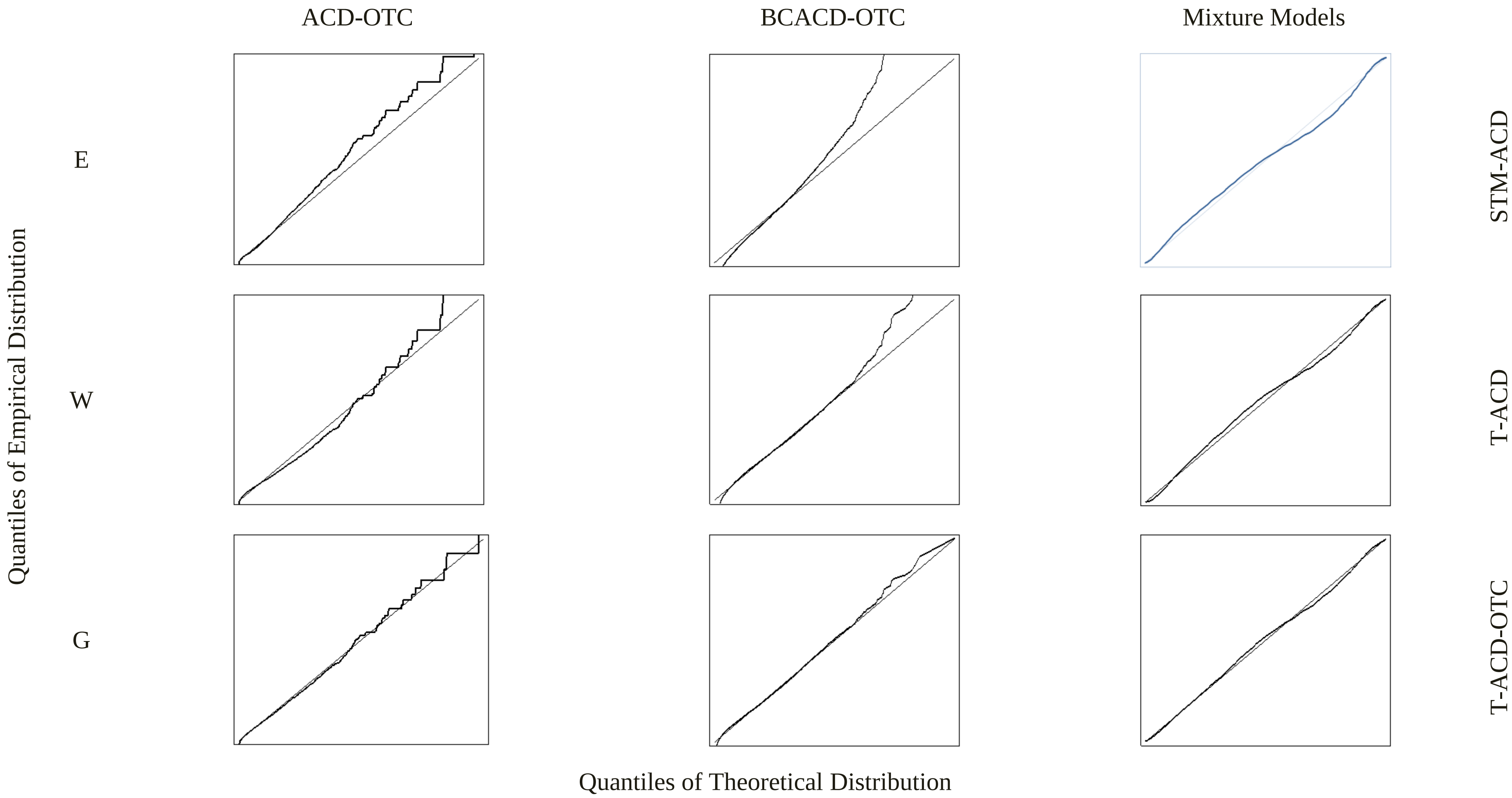
Figure 4.4: Q-Q Plots NP II

A. Existing Models



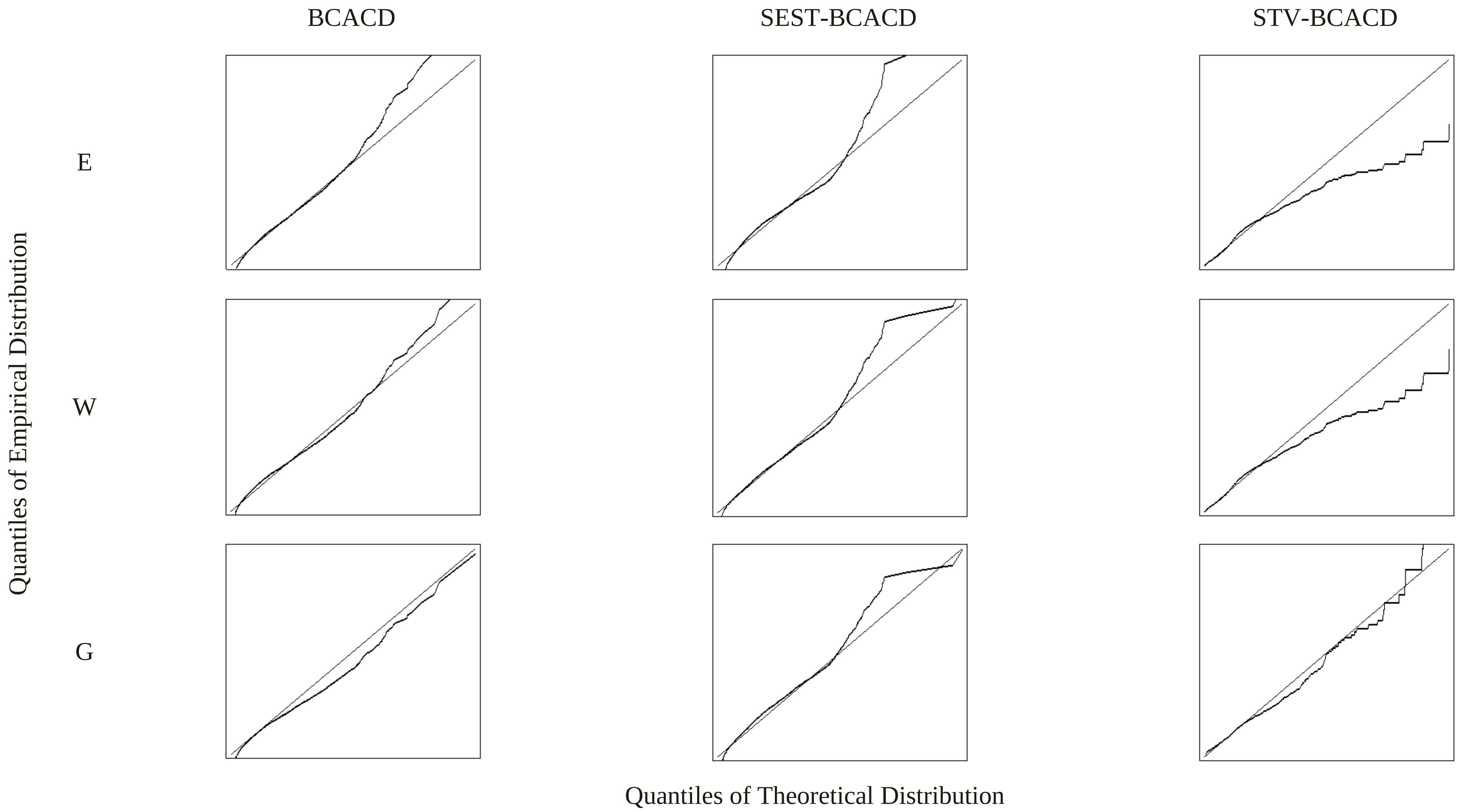
Appendix 4.C

B. ACD-OTC and Mixture of Distributions Models



Appendix 4.C

C. ACD-OTC and Mixture of Distributions Models



Appendix 5.A

Tables

Estimation-Maximum Likelihood

Tables 5.1 (ECX I), 5.2 (ECX II), 5.3 (NP I) and 5.4 (NP II) present the estimation results for all models presented in Section 5.2 in both markets and phases. Each Table is divided horizontally into three parts. In the first part the estimates of all parameters of the STM-ACD model described in Eqs. (5.2), (5.3), (5.4), (5.5), (5.6), (5.7) and (5.8), and of the T-ACD, described in Eqs. (5.10), (5.11), (5.12), (5.13) and (5.14) are presented. The values in parentheses are the associated *t-statistics*. In the second part, *L* stands for the Log-Likelihood function value, *BIC* stands for the Bayesian Information Criterion, while *KS-stats* stands for the Kolmogorov-Smirnov statistics along with critical values. The values in parentheses are the associated *p-values*. The third part presents hypotheses tests for the distribution parameters. The numbers in parentheses are the associated *p-values*.

Table 5.5 presents the estimation results of the Burr-T-ACD models, as in Eqs. (5.15), (5.16) and (5.17). The table is divided into four sections, one for each market in each phase. In each section the first three rows report the estimates of the parameters of the conditional mean specification (i.e., Eq. (5.15)), while the next two refer to the shape parameters of the Burr distribution (i.e., Eqs. (5.16) and (5.17)). Values in parentheses are *t-statistics*. The bottom three rows in each section present hypotheses tests for the distribution parameters. The numbers in parentheses are the associated *p-values*.

Table 5.6 reports the basic statistics (i.e., Mean, Minimum, Maximum and Standard Deviation) of Duration, Volume and Trading Intensity in both markets and phases. The statistics are reported separately for every regime for every year. Furthermore, they are dissected into Buyer (B) and Seller (O) initiated transactions.

Table 5.7 presents the proportion of identical trades and their associated durations in both markets and phases. In more detail, the first three rows of the table present the proportion of total trades that is identical to the previous transaction, when the current trade is normal, fundamental or informed. Along the same lines, the next three rows report the proportion of total trades that is identical, when the current trade is identical to the previous transaction from the same regime (e.g., a current informed transaction is compared to the previous informed). The bottom section of the table shows the average duration for both unique and identical transactions, in both markets and phases, for every regime.

Table 5.8 reports the average duration, volume and trading intensity in both markets and phases, before and after an informed transaction. In more detail, the first panel of the table presents the results for ECX, where the first three rows refer to Phase I and the next three refer to Phase II. In each phase, the results are organized according to the normal, fundamental and informed regimes. The next panel presents the results for NP.

ECX I stands for the data referred to ECX for the Phase I (2005-2007). Respectively, ECX II stands for Phase II (2008) in ECX, NP I for Phase I in Nord Pool and NP II for Phase II of the same market.

Table 5.11: Estimation Results-ECX I

	STM-ACD		Threshold ACD	Threshold ACD-OTC	
Models		Models		Non-OTC	OTC
Coefficients		Coefficients			
omega	0.0258	*omega* $_{normal}$	0.0297	0.0960	0.1027
	(9.88)		(4.42)	(8.75)	(7.44)
alpha	0.1367	*alpha* $_{normal}$	0.0943	0.1130	0.1253
	(5.22)		(6.95)	(12.89)	(14.16)
beta	0.8214	*beta* $_{normal}$	0.8910	0.8120	0.8747
	(33.70)		(25.39)	(44.37)	(48.91)
γ_1	0.9877	*omega* $_{fundamental}$	0.0191	0.0190	0.0370
	(11.40)		(2.42)	(3.20)	(2.56)
γ_2	4.5947	*alpha* $_{fundamental}$	0.1724	0.2427	0.2031
	(12.72)		(4.88)	(9.34)	(4.84)
γ_3	0.3798	*beta* $_{fundamental}$	0.8276	0.7573	0.7969
	(26.91)		(23.44)	(29.16)	(18.99)
g_1	2.3533	*omega* $_{informed}$	0.0021	0.0433	0.0501
	(15.20)		(2.17)	(5.67)	(8.79)
g_2	1.8586	*alpha* $_{informed}$	0.3468	0.2702	0.1355
	(12.84)		(2.93)	(9.28)	(9.70)
s_1	0.3943	*beta* $_{informed}$	0.8614	0.7067	0.8645
	(14.16)		(23.85)	(31.11)	(51.83)
s_2	0.7090				
	(9.65)				
L	-21082.34	*L*	-24802.41057	-23131.36	
BIC	0.9991	*BIC*	1.1668	1.0903	
KS-stat	0.0037		0.0032	0.0030	
$cv_{0.01}=0.0074$	(0.53)		(0.71)	(0.81)	
H(0)					
$\gamma_1=1$	0.04				
	(0.83)				
$\gamma_2=1$	127.46				
	(0.00)				
$\gamma_3=1$	3462.61				
	(0.00)				

Table 5.12: Estimation Results-ECX II

	STM-ACD		Threshold ACD	Threshold ACD-OTC	
Models		Models		Non-OTC	OTC
Coefficients		Coefficients			
omega	0.0702	*omega* $_{normal}$	0.0810	0.0866	0.2442
	(15.62)		(13.04)	(10.26)	(12.51)
alpha	0.0592	*alpha* $_{normal}$	0.1624	0.1685	0.1396
	(68.09)		(22.91)	(20.89)	(15.75)
beta	0.8465	*beta* $_{normal}$	0.7765	0.7543	0.7256
	(49.14)		(69.95)	(29.69)	(36.69)
γ_1	1.0026	*omega* $_{fundamental}$	0.0099	0.0165	0.0186
	(19.91)		(3.49)	(2.12)	(2.87)
γ_2	4.3937	*alpha* $_{fundamental}$	0.2344	0.2653	0.2359
	(21.87)		(9.83)	(7.41)	(9.41)
γ_3	0.6057	*beta* $_{fundamental}$	0.7656	0.7347	0.7641
	(16.98)		(32.11)	(20.51)	(25.75)
g_1	2.3632	*omega* $_{informed}$	0.0085	0.0108	0.0572
	(22.37)		(2.95)	(6.08)	(6.20)
g_2	3.7037	*alpha* $_{informed}$	0.4387	0.1821	0.0547
	(20.73)		(7.09)	(8.35)	(10.78)
s_1	1.0123	*beta* $_{informed}$	0.8385	0.8179	0.8425
	(29.33)		(38.29)	(15.53)	(18.14)
s_2	2.5099				
	(18.52)				
L	60209.48	*L*	-70616.96	-65736.20	
BIC	1.3848	*BIC*	1.5487	1.4430	
KS-stat	0.0019		0.0018	0.0017	
$cv_{0.01}=0.0048$	(0.70)		(0.85)	(0.86)	
H(0)					
$\gamma_1=1$	0.60				
	(0.43)				
$\gamma_2=1$	112.90				
	(0.00)				
$\gamma_3=1$	807.60				
	(0.00)				

Table 5.13: Estimation Results-NP I

	STM-ACD		Threshold ACD	Threshold ACD-OTC	
Models		Models		Non-OTC	OTC
Coefficients		Coefficients			
omega	0.0698	$omega_{normal}$	0.0887	0.1372	0.1319
	(4.53)		(4.21)	(3.34)	(2.82)
alpha	0.2137	$alpha_{normal}$	0.1801	0.1650	0.2039
	(5.84)		(6.29)	(5.92)	(5.85)
beta	0.7013	$beta_{normal}$	0.7500	0.6603	0.7359
	(11.46)		(15.04)	(9.38)	(11.92)
γ_1	1.1226	$omega_{fundamental}$	0.0045	0.0571	0.0478
	(16.55)		(2.25)	(2.05)	(3.45)
γ_2	3.0960	$alpha_{fundamental}$	0.2316	0.4234	0.0093
	(11.77)		(3.24)	(3.72)	(5.08)
γ_3	0.4332	$beta_{fundamental}$	0.7684	0.5766	0.9907
	(21.13)		(10.74)	(6.43)	(8.34)
g_1	2.3840	$omega_{informed}$	0.1387	0.1639	0.1534
	(12.23)		(3.26)	(2.07)	(3.87)
g_2	1.9867	$alpha_{informed}$	0.4138	0.0279	0.3625
	(9.77)		(5.96)	(4.06)	(5.75)
s_1	0.3684	$beta_{informed}$	0.5862	0.4200	0.6375
	(12.61)		(13.61)	(8.44)	(10.84)
s_2	0.9380				
	(8.02)				
L	-1849.76	*L*	-2237.43	-2194.41	
BIC	1.0052	*BIC*	1.1959	1.1949	
KS-stat	0.0120		0.0103	0.0096	
$cv_{0.01}=0.0239$	(0.51)		(0.70)	(0.78)	
H(0)					
$\gamma_1=1$	0.10				
	(0.79)				
$\gamma_2=1$	56.78				
	(0.00)				
$\gamma_3=1$	110.01				
	(0.00)				

Table 5.14: Estimation Results-NP II

	STM-ACD		Threshold ACD	Threshold ACD-OTC	
Models		Models		Non-OTC	OTC
Coefficients		Coefficients			
omega	0.1251	$omega_{normal}$	0.0346	0.0460	0.0040
	(3.48)		(2.73)	(3.25)	(3.40)
alpha	0.1310	$alpha_{normal}$	0.2631	0.1959	0.1483
	(4.39)		(6.79)	(5.91)	(6.19)
beta	0.7625	$beta_{normal}$	0.7369	0.7145	0.8517
	(13.02)		(13.40)	(9.25)	(11.82)
γ_1	1.0629	$omega_{fundamental}$	0.5486	0.4354	0.5478
	(4.00)		(2.98)	(2.41)	(2.80)
γ_2	3.5080	$alpha_{fundamental}$	0.0766	0.0795	0.0786
	(9.79)		(4.76)	(2.09)	(2.93)
γ_3	0.5768	$beta_{fundamental}$	0.4979	0.5228	0.5244
	(12.00)		(5.05)	(2.50)	(3.26)
g_1	1.6797	$omega_{informed}$	0.2923	0.4734	0.5211
	(3.52)		(2.80)	(2.55)	(2.79)
g_2	3.5042	$alpha_{informed}$	0.3995	0.3123	0.1882
	(3.20)		(5.47)	(4.55)	(4.60)
s_1	0.5987	$beta_{informed}$	0.5091	0.3473	0.4242
	(7.87)		(8.53)	(9.23)	(10.80)
s_2	0.9086				
	(4.96)				
L	-2501.60	*L*	-2839.48	-2935.73	
BIC	1.4082	*BIC*	1.5953	1.6714	
KS-stat	0.0104		0.0089	0.0088	
$cv_{0.01}=0.0242$	(0.71)		(0.87)	(0.87)	
H(0)					
$\gamma_1=1$	0.03				
	(0.86)				
$\gamma_2=1$	69.49				
	(0.00)				
$\gamma_3=1$	180.04				
	(0.00)				

Table 5.15: Burr-TACD-All Markets

	ECX I			ECX II		
Models	Normal	Fundamental	Informed	Normal	Fundamental	Informed
Coefficients						
omega	0.0236	0.0513	0.0314	0.1125	0.0448	0.0110
	(3.07)	(4.58)	(10.00)	(9.10)	(6.18)	(2.85)
alpha	0.0634	0.2095	0.7306	0.2717	0.1613	0.1213
	(4.50)	(6.42)	(11.45)	(16.89)	(6.94)	(4.46)
beta	0.9293	0.4505	0.2694	0.6771	0.5352	0.7266
	(34.81)	(4.65)	(7.91)	(34.81)	(7.99)	(8.48)
k	1.0387	1.5937	1.1224	0.9911	2.1229	0.7634
	(9.52)	(12.52)	(14.21)	(23.52)	(17.94)	(14.11)
σ^2	0.0010	(0.39)	1.3540	0.0136	1.1326	0.0064
	(1.91)	(11.35)	(14.69)	(1.12)	(13.29)	(1.32)
H(0)						
k=1	1.27	382.93	1.89	0.45	188.58	88.15
	(0.16)	(0.00)	(0.15)	(0.64)	(0.00)	(0.00)
$\sigma^2=0$	0.37	128.86	165.28	0.14	108.50	0.26
	(0.78)	(0.00)	(0.00)	(0.81)	(0.00)	(0.81)
$\sigma^2=1$	119.27	318.13	113.21	955.69	2.08	157
	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.00)

	NP I			NP II		
Models	Normal	Fundamental	Informed	Normal	Fundamental	Informed
Coefficients						
omega	0.0536	0.1141	0.0069	0.1125	0.0903	0.0097
	(6.99)	(10.61)	(4.23)	(6.20)	(3.41)	(1.65)
alpha	0.1209	0.0076	0.2699	0.1171	0.0515	0.1207
	(9.27)	(2.42)	(4.11)	(9.13)	(3.22)	(2.05)
beta	0.8507	0.6443	0.4493	0.7981	0.6609	0.4019
	(16.67)	(6.15)	(5.97)	(14.29)	(3.87)	(2.42)
k	1.1647	3.0461	1.1628	0.9649	2.9886	1.2289
	(13.33)	(10.17)	(8.41)	(16.01)	(6.16)	(13.60)
σ^2	0.0649	0.3805	0.0596	0.1347	0.4771	0.8104
	(2.16)	(5.98)	(1.76)	(2.26)	(2.95)	3.6372
H(0)						
k=1	1.41	46.64	6.64	4.14	16.82	4.42
	(0.18)	(0.00)	(0.01)	(0.04)	(0.00)	(0.04)
$\sigma^2=0$	2.20	35.76	0.98	6.25	8.67	13.23
	(0.05)	(0.00)	(0.26)	(0.03)	(0.00)	(0.00)
$\sigma^2=1$	21.672	17.11	29.43	46.71	10.41	2.2
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.14)

Table 5.16: Duration, Volume and Trading Intensity: Basic Statistics

			Duration				Volume				Trading Intensity			
ECX			Average	Min	Max	StDev	Average	Min	Max	StDev	Average	Min	Max	StDev
2006	Normal	Total	1640.28	4	30979	3117.01	14.32	1	500	16.68	0.08	0.00	0.40	0.09
		B	1712.82	4	30480	3157.80	15.84	1	500	18.36	0.08	0.00	0.40	0.10
		O	1396.22	4	30979	2963.96	9.19	1	50	6.85	0.07	0.00	0.40	0.09
	Fundamental	Total	67.56	3	1819	108.27	23.72	1	500	31.70	0.54	0.40	0.67	0.08
		B	77.08	3	1819	118.04	27.00	1	500	34.23	0.55	0.40	0.67	0.08
		O	26.78	3	99	15.51	9.66	1	50	7.00	0.53	0.40	0.67	0.08
	Informed	Total	22.83	1	374	40.40	24.07	1	500	40.06	4.13	0.67	78.93	5.83
		B	32.68	1	374	48.49	32.76	1	500	48.76	4.05	0.68	78.93	6.48
		O	7.13	1	76	9.41	10.24	1	50	8.13	4.27	0.67	35.30	4.61
2007	Normal	Total	308.32	3	21091	496.87	9.50	1	500	9.51	0.12	0.00	0.40	0.10
		B	312.19	3	15675	492.90	10.70	1	500	10.49	0.12	0.00	0.40	0.11
		O	301.29	3	21091	503.94	7.31	1	150	6.87	0.10	0.00	0.40	0.10
	Fundamental	Total	44.56	2	840	49.67	15.90	1	250	16.47	0.53	0.40	0.67	0.08
		B	53.67	2	840	54.42	19.04	1	250	18.13	0.53	0.40	0.67	0.08
		O	27.09	2	777	32.50	9.89	1	200	10.28	0.52	0.40	0.67	0.08
	Informed	Total	14.03	1	480	24.64	17.41	1	500	29.61	3.99	0.67	264.25	6.61
		B	19.55	1	480	31.12	22.97	1	500	38.59	3.85	0.67	264.25	7.58
		O	7.77	1	217	11.08	11.09	1	150	10.35	4.15	0.67	68.23	5.28
2008	Normal	Total	92.43	1	4332	170.77	7.42	1	500	11.57	0.24	0.00	1.03	0.27
		B	99.82	1	4332	179.23	9.20	1	500	13.96	0.24	0.00	1.03	0.26
		O	82.47	1	4136	158.11	5.01	1	216	6.42	0.24	0.00	1.03	0.28
	Fundamental	Total	8.63	1	380	14.67	15.15	1	500	26.32	1.62	1.03	2.48	0.40
		B	11.21	1	380	18.47	19.85	1	500	33.36	1.61	1.03	2.48	0.40
		O	5.30	1	114	5.83	9.11	1	190	9.35	1.63	1.03	2.48	0.40
	Informed	Total	3.73	1	141	6.73	23.37	3	500	43.08	6.84	2.48	419.02	8.69
		B	4.72	1	141	8.52	30.38	3	500	54.80	7.27	2.48	419.02	10.22
		O	2.42	1	52	2.42	14.01	3	185	13.20	6.26	2.48	132.11	6.04
NP			Average	Min	Max	StDev	Average	Min	Max	StDev	Average	Min	Max	StDev
2006	Normal	Total	7827.03	68	29933	8037.06	17.50	1	200	22.35	0.04	0.00	0.33	0.06
		B	7979.82	68	29549	8180.61	16.20	1	200	21.38	0.04	0.00	0.33	0.07
		O	7545.57	140	29933	7800.55	19.89	1	120	23.98	0.04	0.00	0.33	0.06
	Fundamental	Total	160.00	37	360	116.98	20.45	5	50	15.57	0.62	0.39	0.81	0.12
		B	156.56	37	360	121.45	21.67	5	50	16.96	0.66	0.54	0.81	0.10
		O	175.50	80	271	135.06	15.00	10	20	7.07	0.47	0.39	0.55	0.11
	Informed	Total	91.64	16	300	82.34	25.45	5	60	16.65	1.53	0.92	2.45	0.44
		B	83.00	16	300	97.52	22.14	5	60	18.45	1.46	0.92	2.11	0.40
		O	106.75	60	180	55.70	31.25	20	50	13.15	1.65	1.21	2.45	0.55
2007	Normal	Total	2121.94	21	28356	3257.22	9.14	1	150	7.67	0.08	0.00	0.34	0.08
		B	2158.97	27	28356	3290.25	9.54	1	100	8.23	0.08	0.00	0.34	0.09
		O	2080.43	21	25982	3220.39	8.69	1	150	6.96	0.07	0.00	0.34	0.08
	Fundamental	Total	109.10	10	2462	137.43	12.46	2	250	14.41	0.55	0.34	0.85	0.15
		B	106.14	10	540	82.83	12.27	2	50	9.34	0.56	0.34	0.85	0.15
		O	112.54	14	2462	181.45	12.67	2	250	18.66	0.54	0.34	0.85	0.14
	Informed	Total	34.38	1	578	50.77	14.22	2	200	16.13	5.92	0.85	120.85	12.06
		B	35.80	1	578	53.17	14.66	2	175	15.51	5.90	0.85	120.85	11.77
		O	32.26	1	450	47.01	13.57	3	200	17.03	5.96	0.86	110.73	12.49
2008	Normal	Total	1608.95	5	20542	2303.49	9.70	1	308	11.26	0.10	0.00	0.68	0.14
		B	1673.41	7	20270	2271.06	10.04	1	215	10.12	0.10	0.00	0.68	0.15
		O	1546.87	5	20542	2333.18	9.37	1	308	12.24	0.10	0.00	0.68	0.14
	Fundamental	Total	95.03	7	1320	176.51	16.37	2	200	27.49	0.77	0.68	0.85	0.05
		B	92.45	7	1320	220.05	16.18	2	200	34.34	0.77	0.68	0.84	0.05
		O	97.61	7	600	121.33	16.55	2	100	18.77	0.77	0.68	0.85	0.05
	Informed	Total	29.80	1	1200	90.44	16.08	1	250	25.13	6.75	0.86	99.49	10.47
		B	27.57	1	616	48.15	16.53	2	200	22.90	7.47	0.88	99.49	12.43
		O	32.01	1	1200	118.32	15.63	1	250	27.19	6.03	0.86	71.92	8.04

Table 5.17: Proportion on Identical Trades and Associated Durations

Previous Trade			ECX I		ECX II		NP I		NP II	
% of Identical Trades	Total	Normal	7.72		8.34		0.00		5.21	
		Fundamental	16.37		9.51		0.00		7.89	
		Informed	31.77		19.51		0.00		29.72	
Same Group			ECX I		ECX II		NP I		NP II	
% of Identical Trades	Total	Normal	8.17		10.08		0.00		4.62	
		Fundamental	6.30		5.31		0.00		0.00	
		Informed	20.94		8.06		0.00		14.29	
			Transactions		Transactions		Transactions		Transactions	
Prevous Trade			Unique	Identical	Unique	Identical	Unique	Identical	Unique	Identical
Average Duration	Total	Normal	552.68	230.89	96.33	49.59	2570.64	N/A	1664.79	594.92
		Fundamental	51.77	19.81	9.02	4.89	110.1811	N/A	98.17	58.33
		Informed	19.00	5.40	3.90	2.11	35.27	N/A	34.09	19.67

Table 5.18: Duration, Volume and Trading Intensity before and after Informed Trading

ECX

		Before			After		
		Duration	Volume	Trading Intensity	Duration	Volume	Trading Intensity
Phase I	Uninformed	1.3504	12.6340	0.8520	1.3526	12.6724	0.8426
	Fundamental	0.8421	12.4998	1.0315	0.8346	12.5955	1.0333
	Informed	0.6708	11.8333	1.4544	0.6613	11.6745	1.4861
Phase II	Uninformed	1.0475	9.5621	0.9213	1.0519	9.6195	0.9301
	Fundamental	0.8705	10.1273	1.1611	0.8527	9.9071	1.1932
	Informed	0.7851	10.3196	1.4237	0.7678	10.0790	1.4799

NP

		Before			After		
		Duration	Volume	Trading Intensity	Duration	Volume	Trading Intensity
Phase I	Uninformed	1.1049	10.9069	0.8588	1.0963	10.8047	0.8888
	Fundamental	0.7581	10.7089	1.0636	0.7282	10.8122	1.1835
	Informed	0.6441	10.3582	1.5232	0.6172	10.6613	1.5754
Phase II	Uninformed	0.9728	10.5877	0.8926	0.9801	10.6648	0.8501
	Fundamental	0.9347	12.0526	0.9940	0.8778	11.1171	1.1734
	Informed	0.7995	10.8739	1.7372	0.7807	10.6502	1.9819

Appendix 5.B

Figures

Empirical Findings

The following Figures provide evidence for the various applications of the estimated models. Figure 5.1 presents the proportion of the different types of trades in both markets and phases and during the trading day, according to the estimation results of STM-ACD (Eqs. (5.5), (5.6) and (5.7)) reported in Tables 5.1, 5.2, 5.3 and 5.4. In more detail, panel A reports the average proportion of normal, fundamental and informed trades in both ECX and NP, both in Phase I and II. Panel B focuses on the proportion of informed trading over the trading years (2006-2008) in both ECX and NP. Panel C presents the proportion of informed trading during the opening, main and closing sessions of both markets in both phases.

Figure 5.2 focuses on the shape parameter of the Weibull distribution, the volume and the autocorrelation of trading. In more detail, panel A of Figure 5.2 graphically presents the estimation results (Tables 5.1, 5.2, 5.3 and 5.4) for the shape parameter of the Weibull distribution associated with each regime in both markets and phases, as well as the transition from one regime to the other. Panel B reports the average trading volume in each regime, while panel C reports the autocorrelation of order flow in ECX and NP and in both phases

Figurer 5.3 presents the average duration (in seconds) of Buyer (B) and Seller (O) initiated transactions in each regime in all markets and phases. Panel A refers to ECX Phase I, while panel B refers to ECX Phase II, panel C refers to NP Phase I and panel D refers to NP Phase II. The first two bars report the average duration in the normal regime, the next two in the fundamental regime and the last two in the informed regime. The actual durations are reported below each bar.

Figure 5.4 focuses on the attributes of OTC transactions. Panel A reports the proportion of OTC transactions in each regime in both markets and phases. Panel B compares the average transaction size (in number of contracts) of normal and OTC transactions.

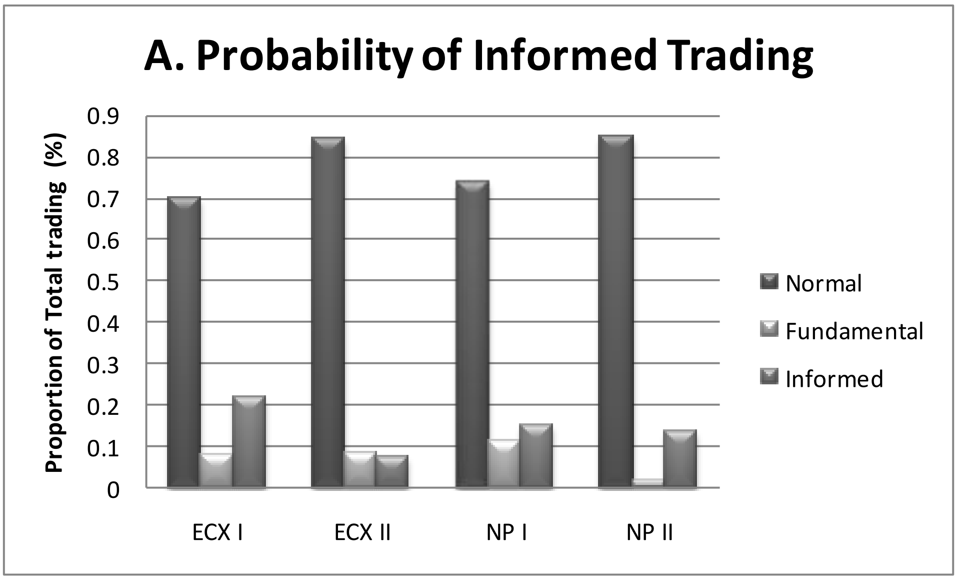
Figure 5.5 reports the relative number of Buyer (Bid) and Seller (Offer) initiated transactions over the years in both markets. In more detail, panel A presents the number of Buys and Sells in ECX, while panel B refers to NP. Panel C reports the difference of Buys and Sells (Buys minus Sells) over the years in both markets.

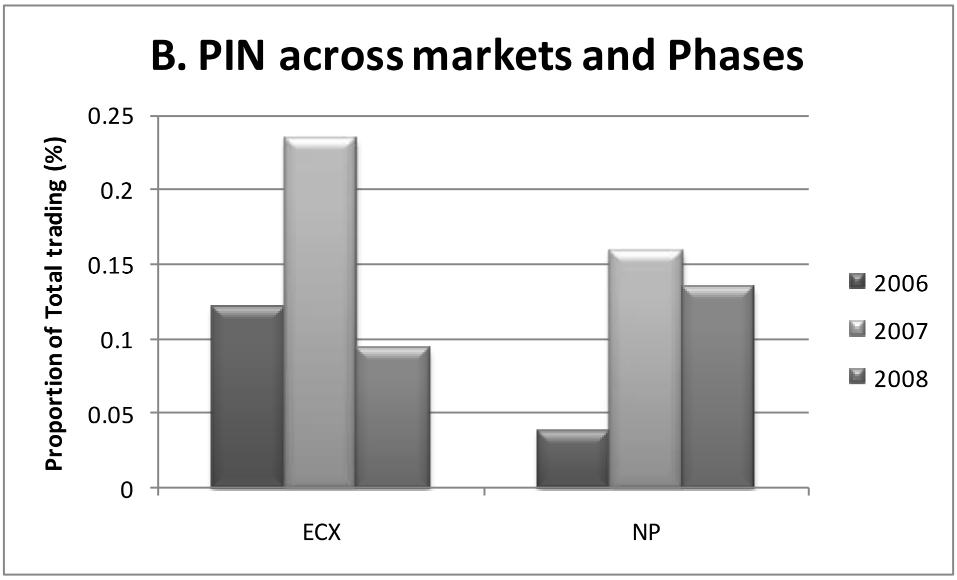
Figures 5.6, 5.7 and 5.8 present various statistics derived from the estimated model (STM-ACD, Eqs. (5.5), (5.6) and (5.7)) around a sharp price drop on 1 August 2008. In more detail, panels A and B present the proportion of normal, fundamental and informed trading in ECX and NP, respectively. Figure 5.7 reports the difference in the number of Buyer and Seller (Buyer minus Seller) initiated transactions, dissected to the different regimes around the price drop. Panel A refers to ECX, while panel B refers to NP. Along the same lines, Figure 5.8 presents the relative volumes and durations of Buys and Sells in each regime (normal, fundamental and informed) in ECX (panel A) and NP (panel B), three days before and three days after the price drop.

Finally, Figure 5.9 presents the hazard functions derived from the estimation results of Burr-T-ACD model, as in Eqs. (5.15), (5.16) and (5.17), reported in Table 5.5. The first column presents the hazard function of the durations associated with low trading intensity (normal regime) in both markets and phases. Similarly, the second column presents the hazard functions of the durations associated with the fundamental regime, while the last column refers to the informed regime.

Appendix 5.B

Figure 5.1: Probability of Informed Trading (PIN)





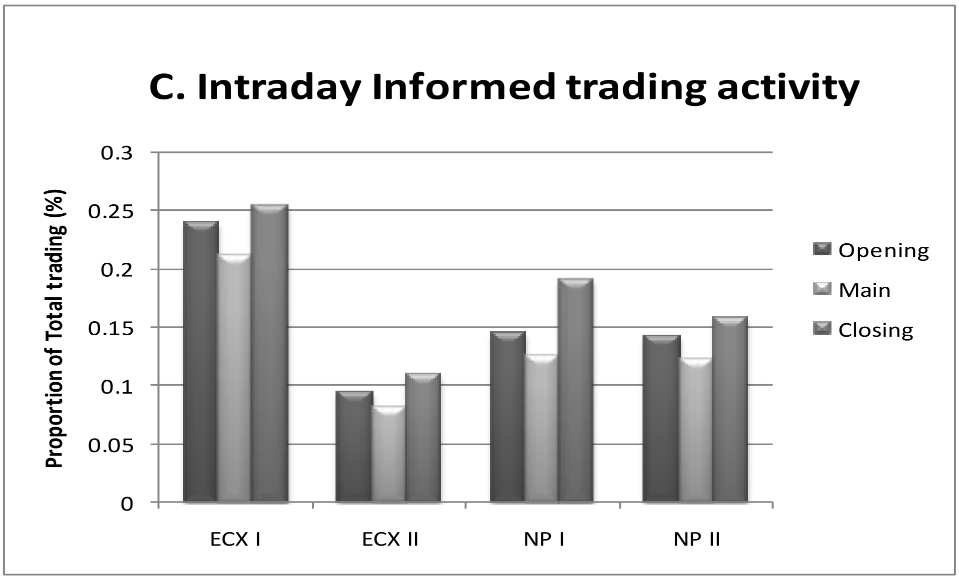
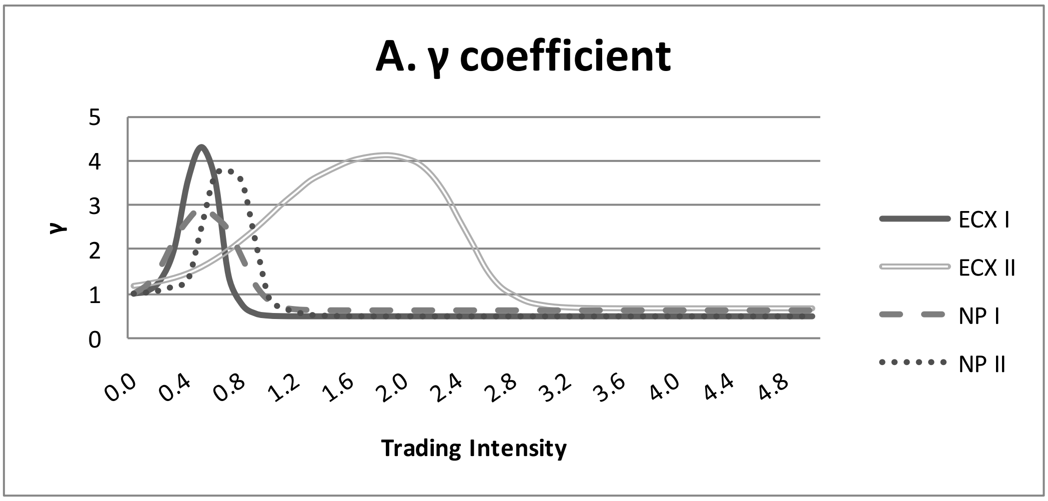
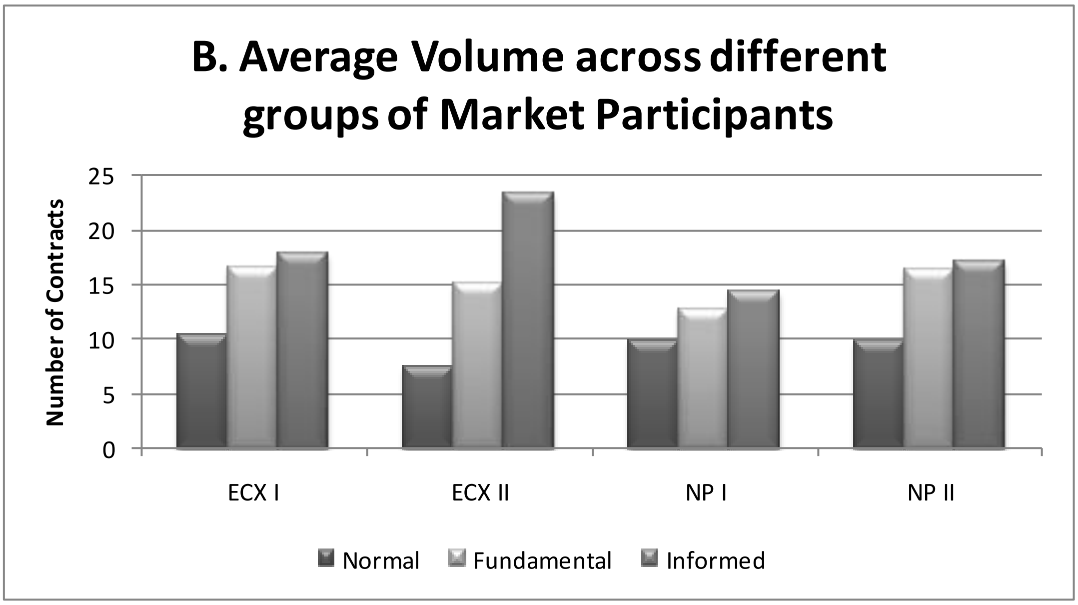
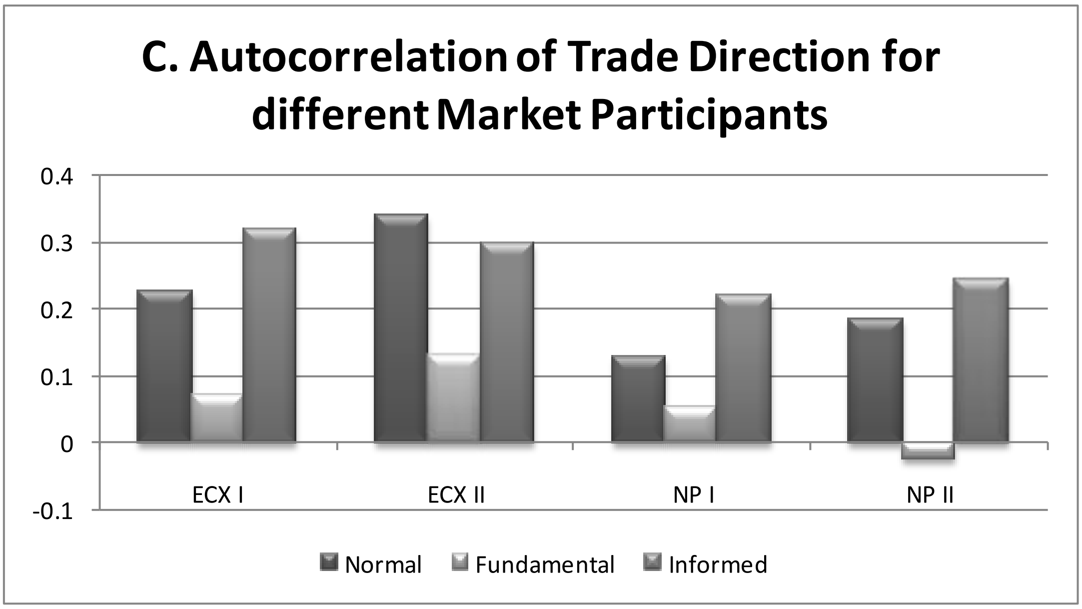


Figure 5.2: Average Volume and Autocorrelations of Regime Trades

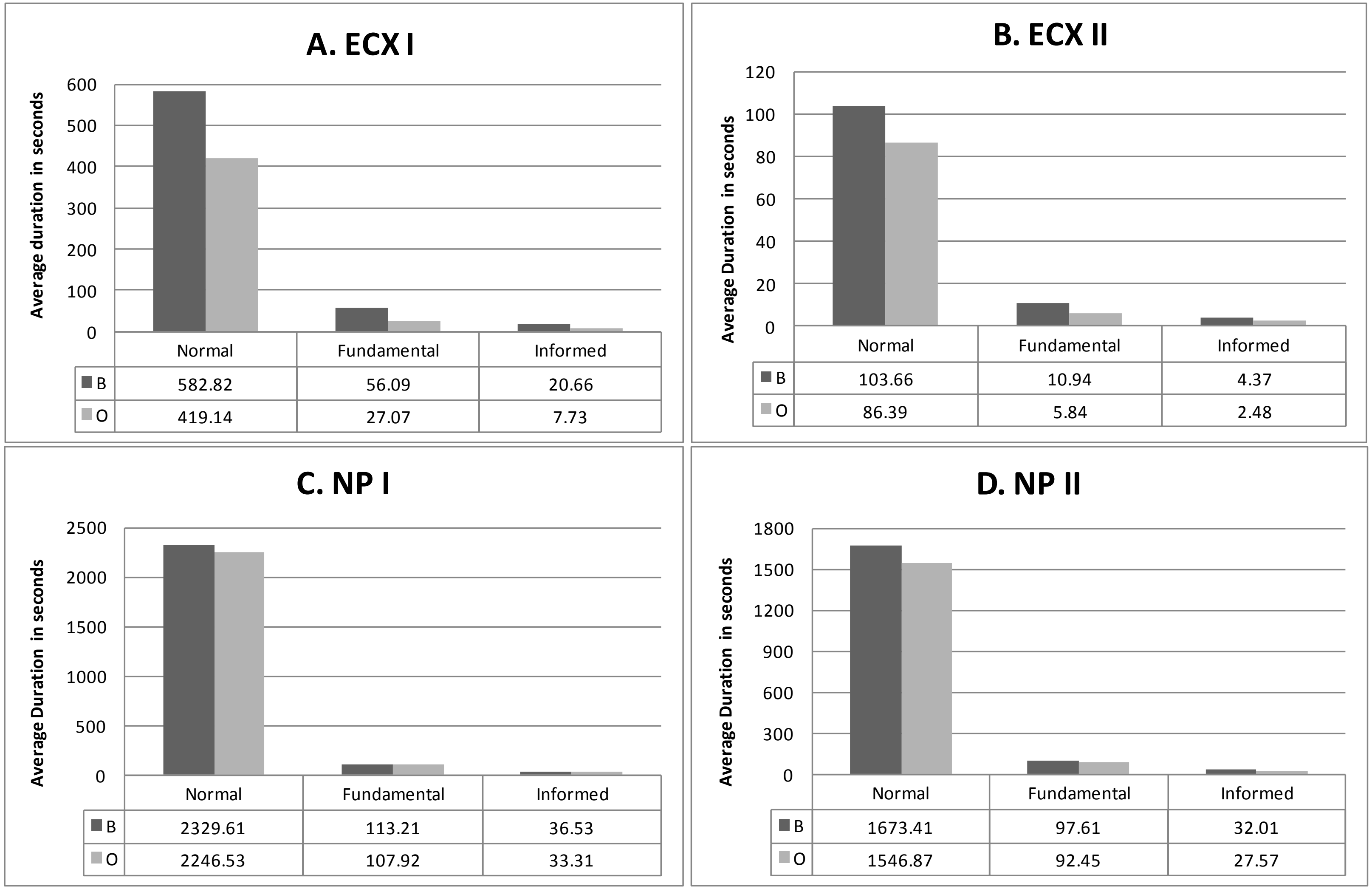






Appendix 5.B

Figure 5.3: Duration by Market, Trader Type and Trade Sign Indicator



A. ECX I

	Normal	Fundamental	Informed
B	582.82	56.09	20.66
O	419.14	27.07	7.73

B. ECX II

	Normal	Fundamental	Informed
B	103.66	10.94	4.37
O	86.39	5.84	2.48

C. NP I

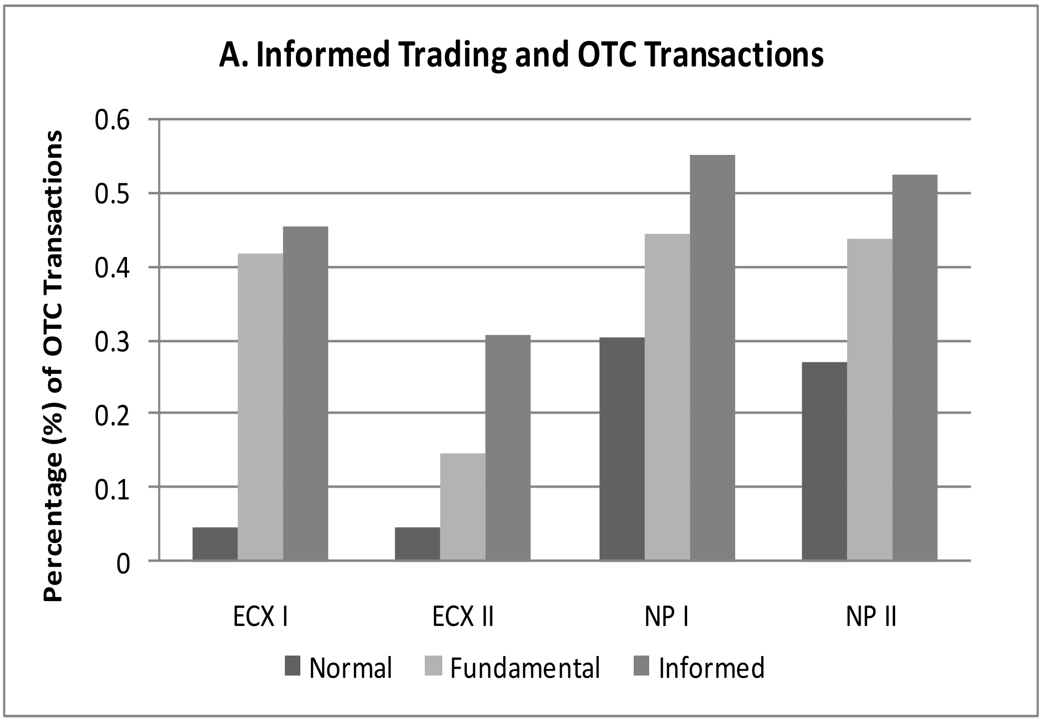
	Normal	Fundamental	Informed
B	2329.61	113.21	36.53
O	2246.53	107.92	33.31

D. NP II

	Normal	Fundamental	Informed
B	1673.41	97.61	32.01
O	1546.87	92.45	27.57

Appendix 5.B

Figure 5.4: Average Duration and Proportion of OTC Trades



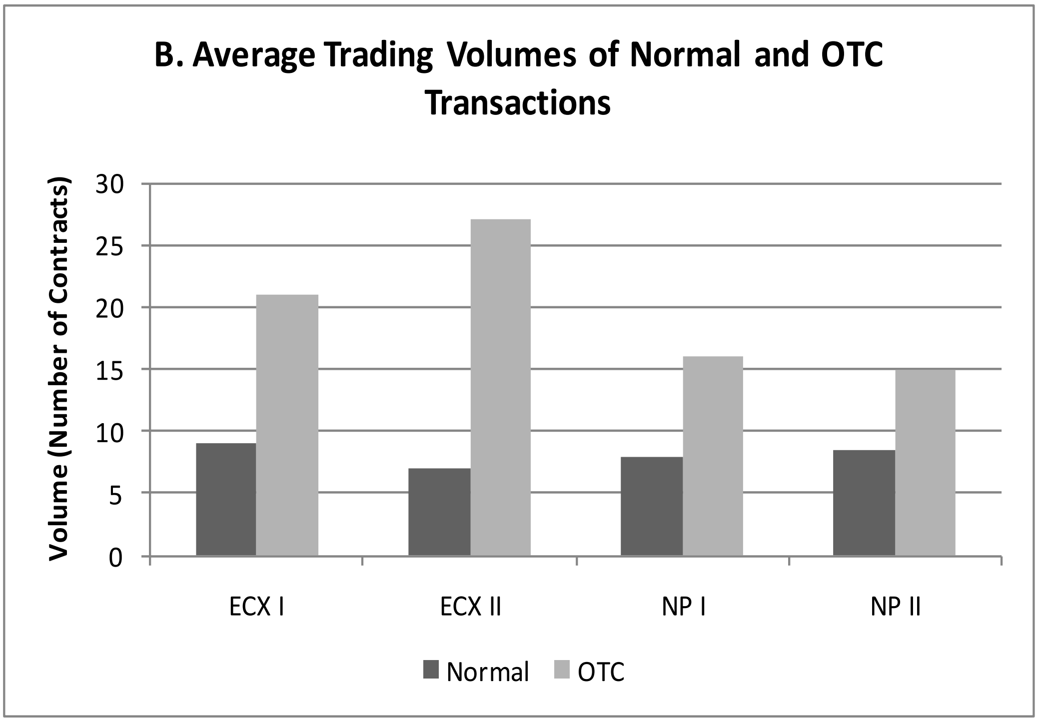


Figure 5.5: Volume by Trade Sign Indicator

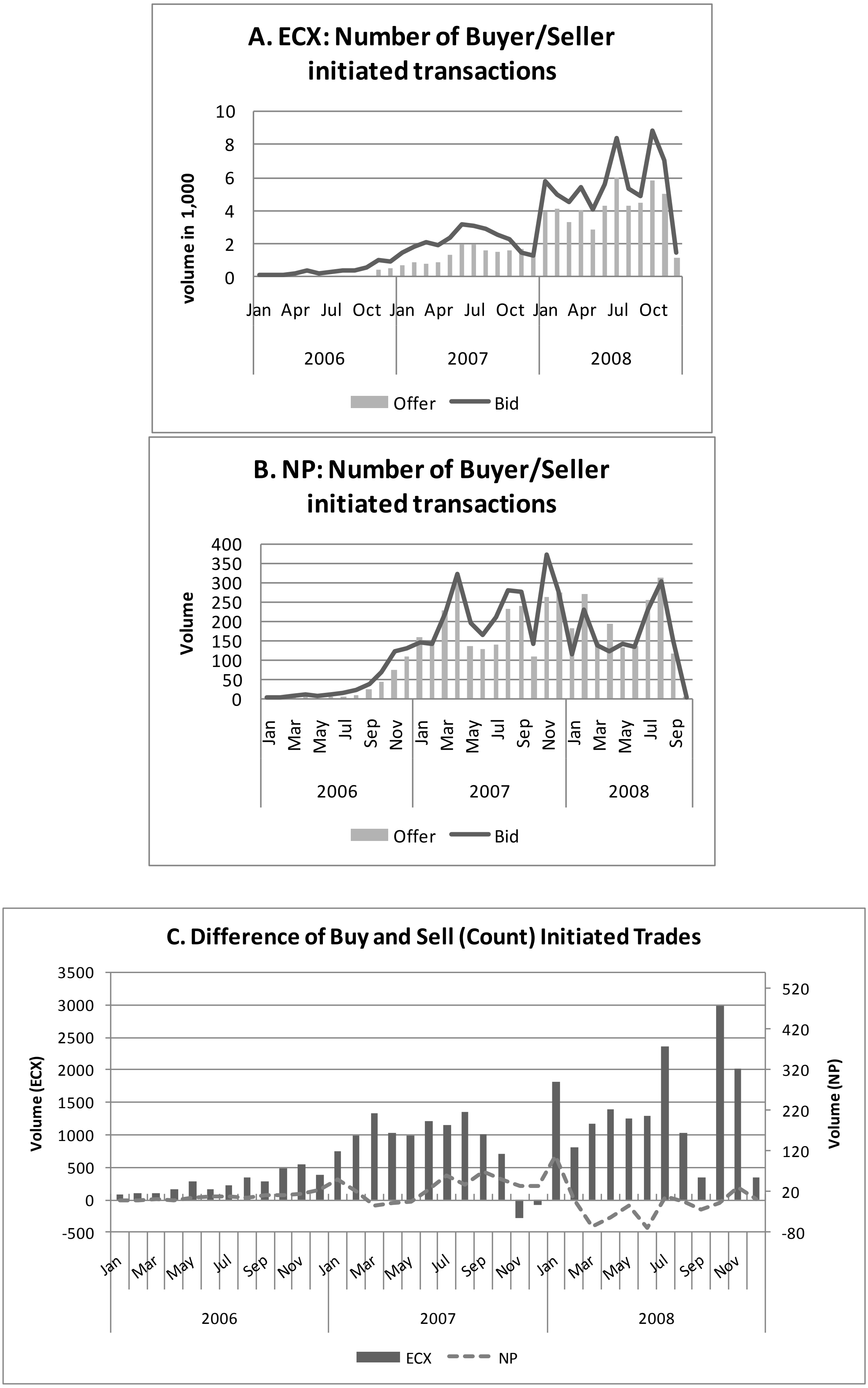
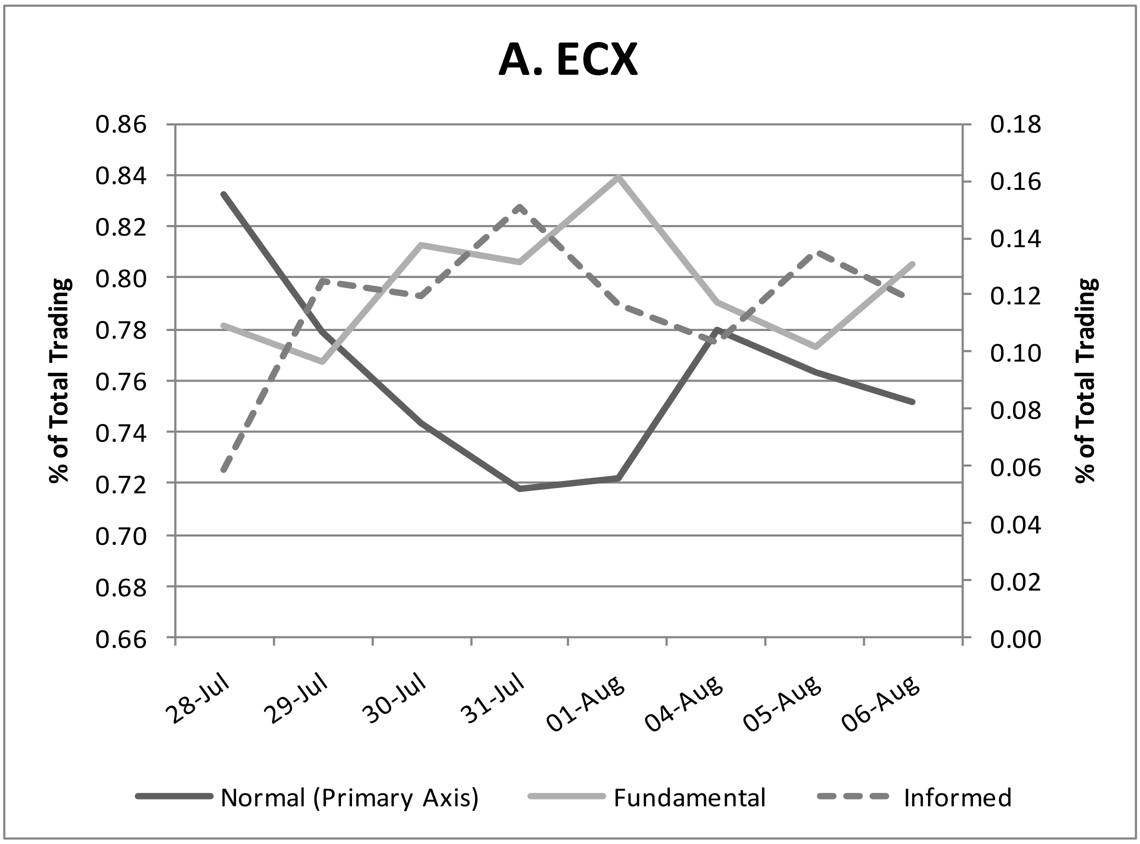


Figure 5.6: Proportion of Trader Type Around a Negative Price Event



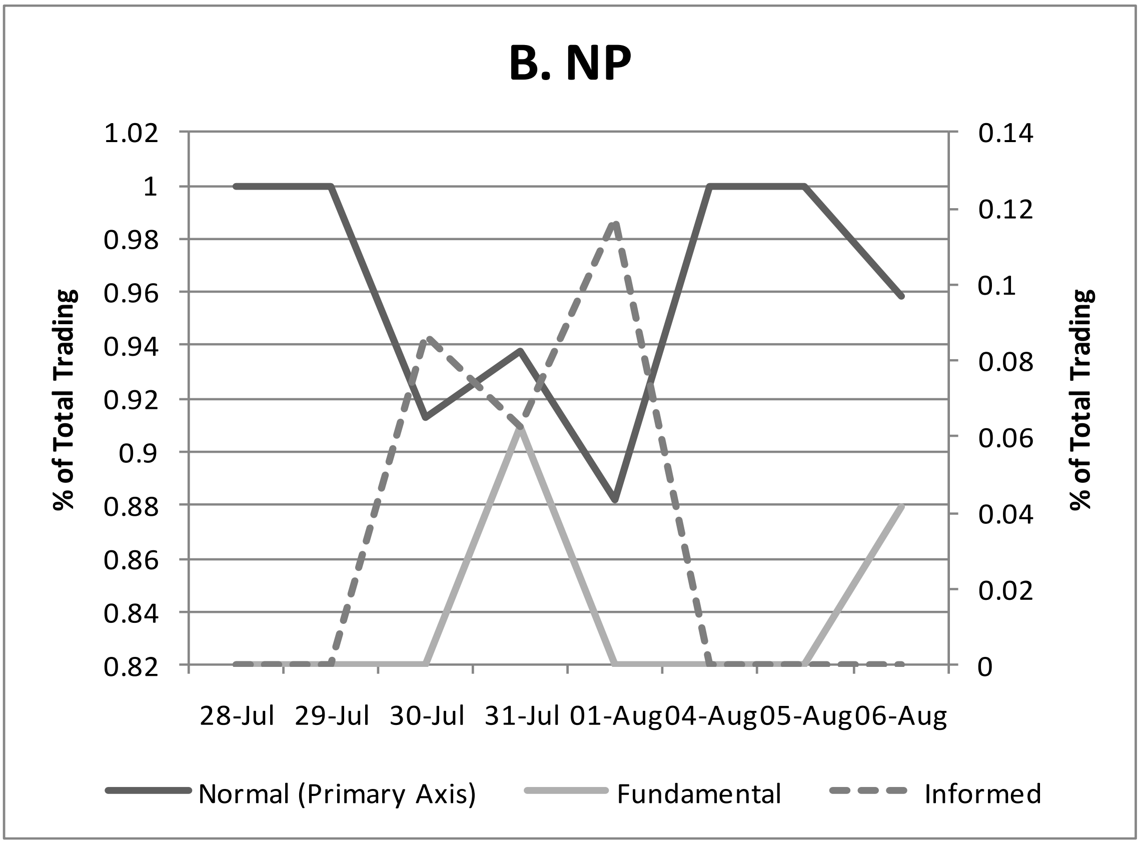
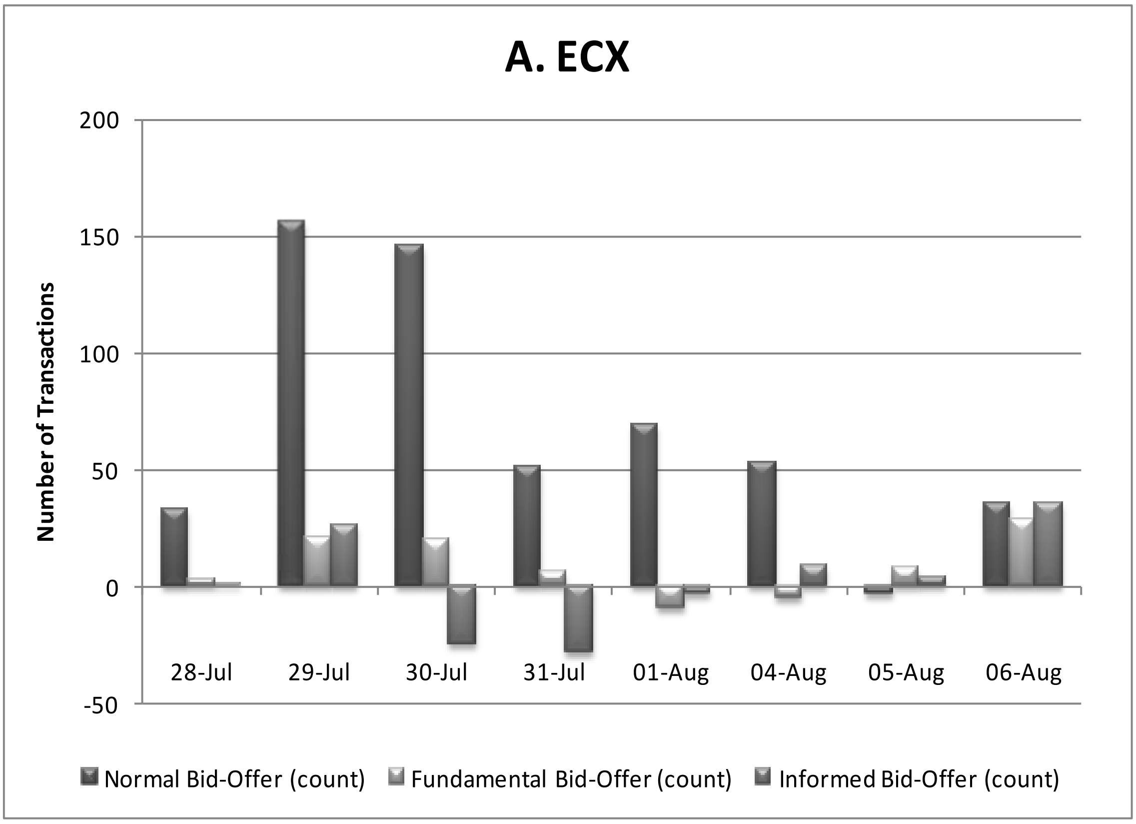


Figure 5.5: Trades by Sign (Difference between Bid and Offer) Around a Negative Event



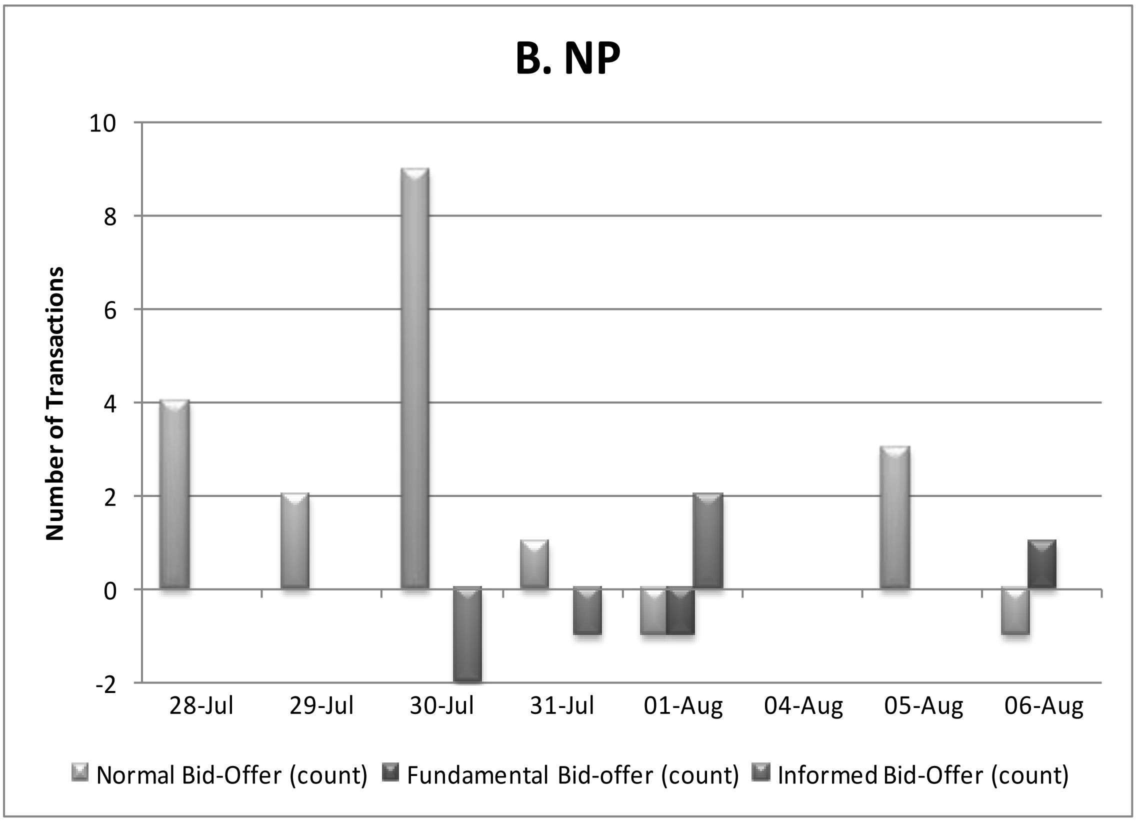
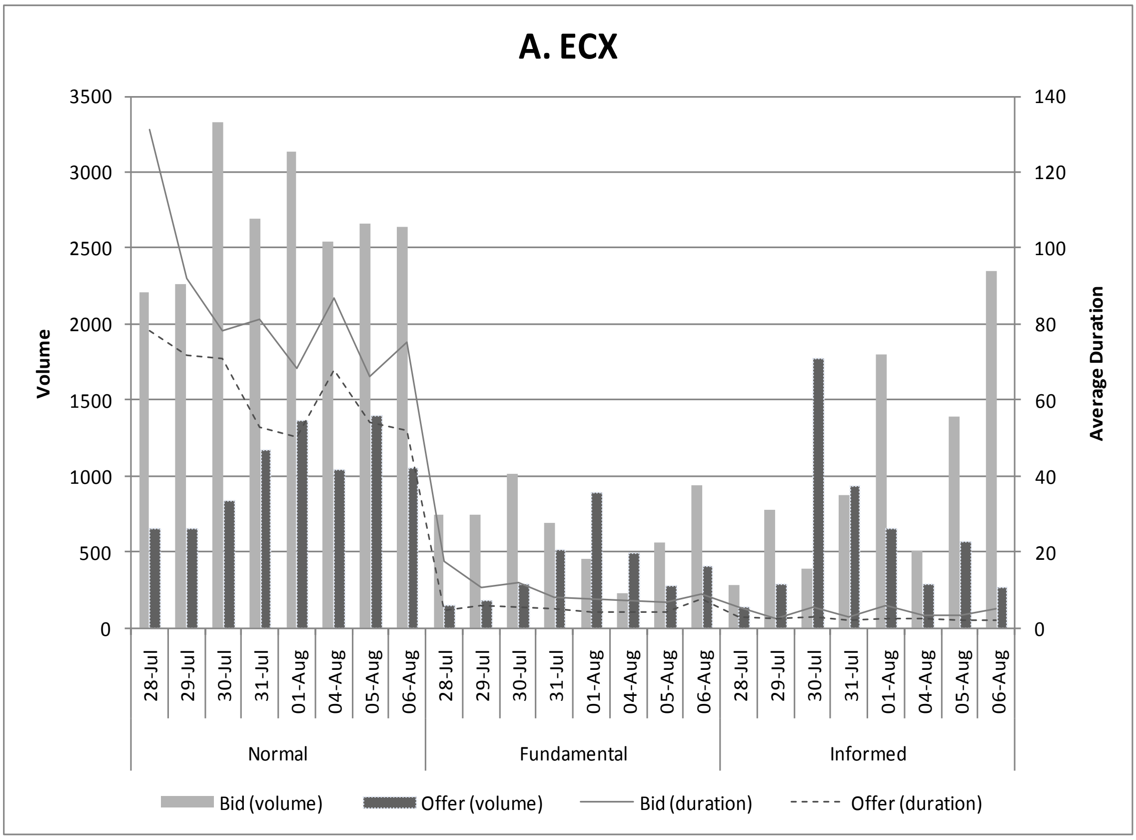


Figure 5.8: Sum of Signed Volume and Average Duration by Trader Type Around a Negative Price Event



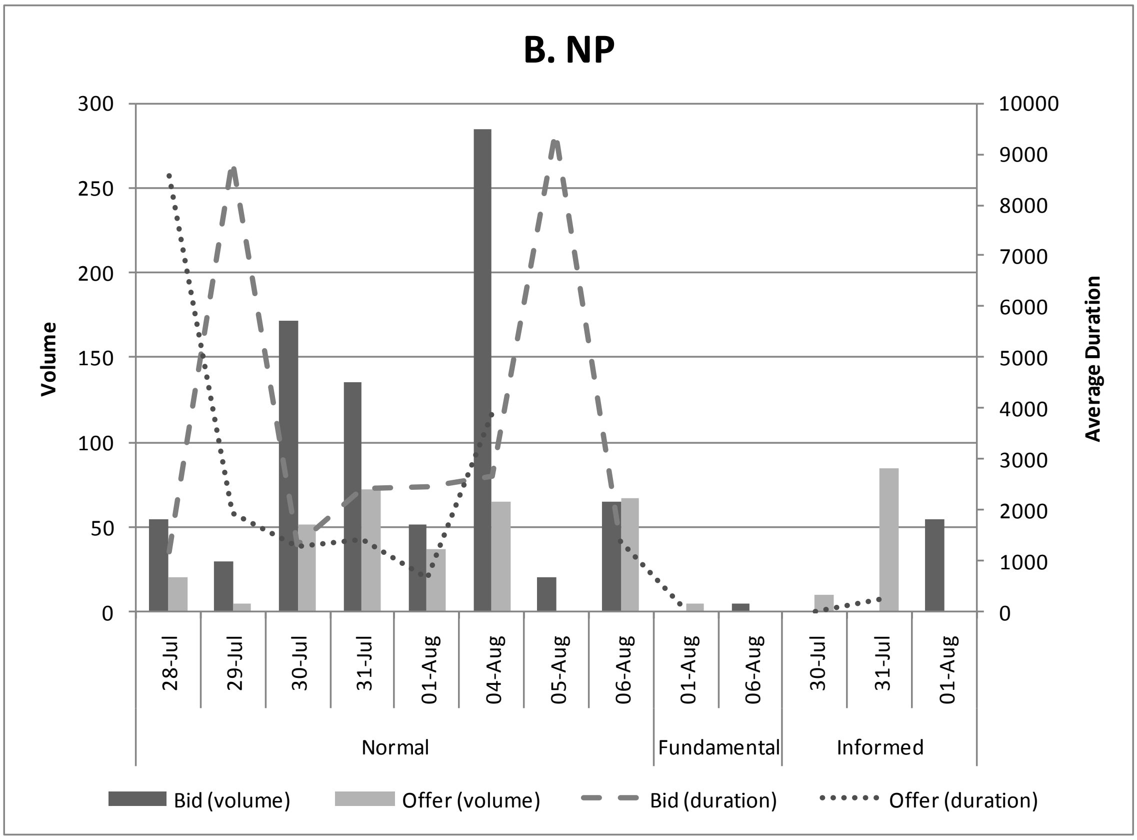
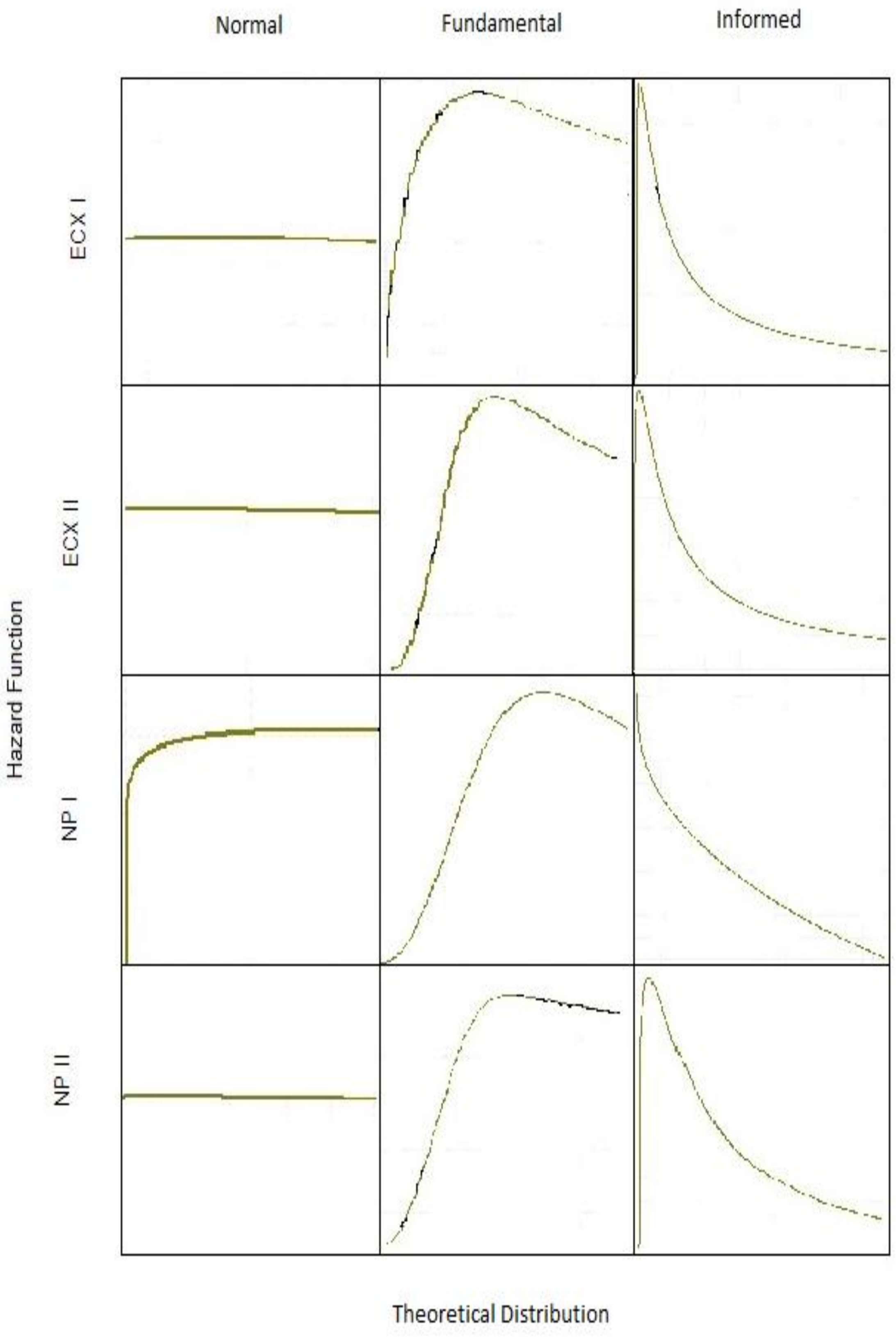


Figure 5.9: Hazard functions For Burr-TACD, All Markets



Appendix 6.A

Derivation of Equations

The next section presents the derivation of model equations required for post estimation analysis. More specifically, in the first part, the second moment of price change, Δp_t, (Eq. (6.27) in Section 6.2.2) is analyzed taking into account all relevant model assumptions. The second part explicitly describes the derivation of the partial derivative of price change with respects to expected trading intensity, $E[s_t|H_{t-1}]$, discussed in Eq. (6.31) in Section 6.2.2.

Eq. (6.27) in Section 6.2.2

The derivation of the second moment of the price change can be expressed as:

$$E[\Delta p_t]^2 = E\left[\underbrace{\left(\varepsilon_t + (\xi_t - \xi_{t-1})\right)}_{a} + \underbrace{(\theta_t + \varphi_t)q_t - (\rho\theta_t + \varphi_{t-1})q_{t-1}}_{b}\right]^2$$

$$= \underbrace{E\varepsilon_t{}^2}_{\sigma_\varepsilon^2} + \underbrace{E(\xi_t - \xi_{t-1})^2}_{=2\sigma_\xi^2} + \underbrace{E2\varepsilon_t(\xi_t - \xi_{t-1})}_{=0} + E\{b\}^2 + \underbrace{E2ab}_{=0}$$

$$= \sigma_\varepsilon^2 + 2\sigma_\xi^2 + \{(\theta_t + \varphi_t)q_t - (\rho\theta_t + \varphi_{t-1})q_{t-1}\}^2$$

$$= \sigma_\varepsilon^2 + 2\sigma_\xi^2 + \{(\theta_t + \varphi_t)^2 + (\rho\theta_t + \varphi_{t-1})^2 - 2\rho(\theta_t + \varphi_t)(\rho\theta_t + \varphi_{t-1})\}.$$

This variance Expression is evaluated under $E\varepsilon_t = 0, \mathrm{E}\xi_t = 0, E\varepsilon_t{}^2 = \sigma_\varepsilon^2, E\xi_t{}^2 = 2\sigma_\xi^2$, $Eq_t = 0, \mathrm{E}q_t{}^2 = 1, \mathrm{E}q_t q_{t-1} = \rho$.

Eq. (6.31) in Section 6.2.2

The derivation of the first order partial derivative of the price change variance, with respect to $E[s_t|H_{t-1}]$:

The return variance can be written as:

$$\underbrace{Z}_{Var(\Delta p_t)} = \underbrace{a}_{\sigma_\varepsilon^2 + 2\sigma_\xi^2} + \underbrace{[f(\theta_t, \varphi_t)]^2}_{(\theta_t + \varphi_t)^2} + \underbrace{[g(\theta_t, \varphi_{t-1})]^2}_{(\rho * \theta_t + \varphi_{t-1})^2} - 2\rho \underbrace{f(\theta_t, \varphi_t)}_{(\theta_t + \varphi_t)} \underbrace{g(\theta_t, \varphi_{t-1})}_{\rho * \theta_t + \varphi_{t-1}}.$$

Then, applying the following rules of partial differentiation:

- The generalized power function rule:

$$\frac{\partial[f(\theta_t, \varphi_t)]^2}{\partial E[s_t|H_{t-1}]} = 2(\theta_t + \varphi_t)\frac{\partial(\theta_t + \varphi_t)}{\partial E[s_t|H_{t-1}]}$$

$$= 2(\theta_t + \varphi_t)\frac{\partial\left\{\theta\left(s, \tilde{I}, \sigma_p^2, \psi\right) + \varphi\left(s, \tilde{I}, \sigma_p^2, \psi\right)\right\}}{\partial E[s_t|H_{t-1}]}$$

$$= 2(\theta_t + \varphi_t)\left\{\sum_{\pi}^{3}\left[\left(\theta_{2,\pi} + \varphi_{2,\pi}\right)I_{\pi,t}\right]\right\},$$

and

$$\frac{\partial[g(\theta_t,\varphi_{t-1})]^2}{\partial E[s_t|H_{t-1}]} = 2(\rho\theta_t + \varphi_{t-1})\frac{\partial(\rho\theta_t + \varphi_{t-1})}{\partial E[s_t|H_{t-1}]}$$

$$= 2(\rho\theta_t + \varphi_{t-1})\frac{\partial\big(\rho\theta\big(s\,,\tilde{I}\,,\sigma_p^2,\psi\,\big) + \varphi_{t-1}\big)}{\partial E[s_t|H_{t-1}]}$$

$$= 2\rho(\rho\theta_t + \varphi_{t-1})\left\{\sum_{\pi}^{3}\big(\theta_{2,\pi} I_{\pi,t}\big)\right\}.$$

- The product of functions rule:

$$\frac{\partial[2\rho f(\theta_t,\varphi_t)g(\theta_t,\varphi_{t-1})]}{\partial E[s_t|H_{t-1}]} =$$

$$= 2\rho\left\{\frac{\partial[f(\theta_t,\varphi_t)]}{\partial E[s_t|H_{t-1}]}g(\theta_t,\varphi_{t-1}) + \frac{\partial[g(\theta_t,\varphi_{t-1})]}{\partial E[s_t|H_{t-1}]}f(\theta_t,\varphi_t)\right\}$$

$$= 2\rho\left\{\frac{\partial\big\{\theta\big(s\,,\tilde{I}\,,\sigma_p^2,\psi\,\big) + \varphi\big(s\,,\tilde{I}\,,\sigma_p^2,\psi\,\big)\big\}}{\partial E[s_t|H_{t-1}]}(\rho\theta_t + \varphi_{t-1})\right.$$

$$\left. + \frac{\partial\big(\rho\theta\big(s\,,\tilde{I}\,,\sigma_p^2,\psi\,\big) + \varphi_{t-1}\big)}{\partial E[s_t|H_{t-1}]}(\theta_t + \varphi_t)\right\}$$

$$= 2\rho\left\{(\rho\theta_t + \varphi_{t-1})\left\{\sum_{\pi}^{3}\big[\big(\theta_{2,\pi} + \varphi_{2,\pi}\big)I_{\pi,t}\big]\right\} + (\theta_t + \varphi_t)\left\{\sum_{\pi}^{3}\big(\theta_{2,\pi} I_{\pi,t}\big)\right\}\right\}.$$

Then, the first order partial derivative of $Var(\Delta p_t)$, with respect to $E[s_t|H_{t-1}]$, after gathering like terms can be written as:

$$\frac{\partial[Var(\Delta p_t)]}{\partial E[s_t|H_{t-1}]} = 2\left\{(\theta_t + \varphi_t)\left\{\sum_{\pi}^{3}\big[\big(\theta_{2,\pi}(1-\rho^2) + \varphi_{2,\pi}\big)I_{\pi,t}\big]\right\}\right.$$

$$\left. - (\rho\theta_t + \varphi_{t-1})\left\{\sum_{\pi}^{3}\big(\varphi_{2,\pi} I_{\pi,t}\big)\right\}\right\}.$$

Appendix 6.B

Tables

Estimation Generalized Method of Moments (GMM)

and Post Estimation Analysis

The following section presents the estimation results for the MRR and the Dynamic Expectations Joint Model, discussed in Section 6.2, as well as a tabulated post estimation analysis of price changes, spreads, price change variance and their components, based on basic statistics.

More specifically, Table 6.1 presents the estimates of the coefficients of Eqs. (6.22) and (6.23), with moment conditions in Eq. (6.28). *t-statistics* are in parentheses. The first column presents the original, MRR, model, while the next three columns examine explicitly the information spread component, θ_t, applying a stepwise regression analysis, incorporating trading intensity, T.I. and the Risk, as in Eq. (6.16). The next three columns examine in a similar fashion the liquidity spread component, φ_t. The last two columns present the estimation results of the joint model, where in the last column only the statistically significant parameters are included. The coefficients are arranged as follows. First, θ_t (Eq. (6.16)) is analyzed, being followed by φ_t (Eq. (6.18)), the autocorrelation of the order flow, ρ, and the estimates of the coefficients in Eq. (6.12). The bottom two sections present the *J-statistics* and the associated *p-values* in parentheses, as well as the Mean Squared Error for in- and out-of-sample forecasts.

Table 6.2 presents the average trading intensity (arithmetic mean) and the parameter estimates of the joint model presented in Eqs. (6.22) and (6.23), as well their product, for every regime of trading intensity, as they have been found in Chapter 5. The last column is the implied (Eq. (6.24)) half spread (excluding the risk aversion component).

Appendix 6.B

Table 6.3 presents the intraday variations of Actual and Expected Trading Intensity, Estimated Spread and Variance and their Components.

Table 6.4.A presents the average implied spread, as in Eq. (6.24), and adverse-selection component, θ_t, as in Eq. (6.16), across different levels of transaction size and frequency, in all markets and phases. The threshold values have been arbitrarily chosen to indicate low, middle and high trading activity and they are the same for both markets in both phases. Furthermore, the table differentiates the relative sizes of spreads between Buyer and Seller initiated transactions. Focusing on each market, the first two columns present the width of the implied spread and the size of the adverse-selection component after a trade that belongs to one of the pre-described different regimes. The next two columns summarize the values of spreads and adverse-selection across different levels of trading intensity of the current transaction. Table 6.4.B presents the relative values for the second moment of price change.

Finally, the abbreviations ECX I, ECX II, NP I and NP II stand for European Climate Exchange Phase I, European Climate Exchange Phase II, Nord Pool Phase I and Nord Pool Phase II, respectively.

Table 6.19: Estimation Results

A. ECX I

		Theta			Phi			Full	
	MRR	T.I	Risk	T.I+Risk	T.I	Risk	T.I+Risk	All	Significant
ϑ_1	0.0808	0.0771	0.0762	0.0707	0.0816	0.0804	0.0821	0.0112	
	(9.31)	(2.06)	(5.94)	(2.03)	(9.05)	(8.68)	(9.13)	(0.35)	
$\vartheta_{2,\ uninformed}$		0.0159		0.0186				0.0921	0.0987
		(1.62)		(1.71)				(2.23)	(5.45)
$\vartheta_{2,\ fundamental}$		0.0077		0.0095				0.0698	0.0469
		(2.42)		(2.51)				(2.36)	(3.04)
$\vartheta_{2,\ informed}$		0.0017		0.0016				0.0206	0.0398
		(2.97)		(2.79)				(3.84)	(2.93)
ϑ_3			0.0269	0.0307				0.0756	
			(0.43)	(0.47)				(0.88)	
ϑ_4			0.0004	0.0004				-0.0005	
			(0.66)	(0.68)				(-0.76)	
φ_1	0.0410	0.0393	0.0413	0.0391	0.0760	0.0317	0.0701	0.0985	0.0969
	(3.26)	(3.31)	(3.02)	(3.32)	(3.27)	(2.71)	(2.45)	(3.33)	(3.40)
$\varphi_{2,\ uninformed}$					-0.0337		-0.0313	-0.0834	-0.0889
					(-1.96)		(-1.87)	(-2.19)	(-2.24)
$\varphi_{2,\ fundamental}$					-0.0261		-0.0213	-0.0561	-0.0539
					(-2.39)		(-2.21)	(-2.87)	(-2.84)
$\varphi_{2,\ informed}$					-0.0124		-0.0125	-0.0124	-0.0127
					(-3.19)		(-2.98)	(3.39)	(2.96)
φ_3						0.0454	0.0060	-0.0457	
						(0.58)	(0.75)	(-0.43)	
$\varphi 4$						0.0011	0.0010	0.0015	0.0011
						(2.58)	(2.48)	(2.64)	(2.64)
p_1	0.5022	0.5002	0.5001	0.5007	0.5003	0.5003	0.5000	0.4999	0.5001
	(17.08)	(17.20)	(17.16)	(17.22)	(17.09)	(17.15)	(17.08)	(17.07)	(17.11)
c_1		0.5965		0.5971	0.5960		0.5961	0.5945	0.5963
		(14.00)		(14.02)	(13.97)		(13.97)	(13.90)	(14.08)
c_2		0.2626		0.2631	0.2637		0.2638	0.2641	0.2632
		(15.25)		(15.28)	(15.26)		(15.27)	(15.25)	(15.24)
c_3		0.0010		0.0002	-0.0004		-0.0005	-0.0016	
		(0.09)		(0.09)	(-0.25)		(-0.45)	(-0.14)	
c_4		0.0128		0.0129	0.0171		0.0171	0.0181	
		(0.66)		(0.67)	(0.87)		(0.87)	(0.92)	
c_5		0.0354		0.0346	0.0374		0.0374	0.0372	
		(0.39)		(0.69)	(0.63)		(0.73)	(0.72)	
c_6		-1.1828		-1.1828	-1.1879		-1.1880	-1.1927	-1.2069
		(-24.41)		(-24.40)	(-24.24)		(-24.20)	(-24.21)	(-28.78)
σ_ε^2	0.0066	0.0082	0.0082	0.0082	0.0082	0.0082	0.0082	0.0082	0.0082
	(2.24)	(2.11)	(2.04)	(1.96)	(1.99)	(2.02)	(1.97)	(2.01)	(2.01)
σ_ξ^2	0.00017	0.00020	0.00020	0.00020	0.00020	0.00020	0.00020	0.00020	0.00020
	(2.32)	(2.31)	(2.18)	(2.24)	(2.23)	(2.21)	(2.26)	(2.22)	(2.29)
J-stats	2.50	8.71	7.65	9.86	8.13	6.47	11.70	2.52	1.57
	(0.11)	(0.07)	(0.05)	(0.13)	(0.08)	(0.06)	(0.07)	(0.11)	(0.21)
$MSE_{in\ sample}$	1.5058	1.5033	1.6494	1.5013	1.4996	1.5075	1.4990	1.4981	1.4945
$MSE_{out\text{-}of\text{-}sample}$	2.3101	2.3065	2.4662	2.3027	2.2965	2.3023	2.2961	2.2957	2.2946

B. ECX II

		Theta			Phi			Full	
	MRR	T.I	Risk	T.I+Risk	T.I	Risk	T.I+Risk	All	Significant
ϑ_1	0.0496	0.0326	0.0478	0.0272	0.0498	0.0497	0.0498	-0.0090	
	(15.38)	(2.02)	(10.86)	(1.75)	(15.51)	(15.25)	(15.52)	(-0.40)	
$\vartheta 1_{, uninformed}$		0.0242		0.0366				0.0715	0.0613
		(2.27)		(2.39)				(2.39)	(8.31)
$\vartheta 1_{, fundamental}$		0.0133		0.0150				0.0456	0.0391
		(2.31)		(2.47)				(2.98)	(9.85)
$\vartheta 1_{, informed}$		0.0044		0.0049				0.0248	0.0286
		(3.50)		(3.57)				(3.07)	(9.91)
$\vartheta 2$			0.0313	0.0473				0.0383	
			(0.91)	(1.35)				(0.94)	
$\vartheta 3$			-0.0003	-0.0003				-0.0014	
			(-0.04)	(-0.34)				(-1.83)	
φ_1	0.0182	0.0126	0.0165	0.0117	0.0308	0.0083	0.0159	0.0590	0.0599
	(3.10)	(2.35)	(2.85)	(2.20)	(2.94)	(2.07)	(1.93)	(2.44)	(3.76)
$\varphi_{2, uninformed}$					-0.0117		-0.0068	-0.0561	-0.0522
					(-2.54)		(-2.11)	(-2.02)	(-2.63)
$\varphi_{2, fundamental}$					-0.0109		-0.0042	-0.0401	-0.0385
					(-2.93)		(-2.76)	(-2.61)	(-3.57)
$\varphi_{2, informed}$					-0.0035		-0.0011	-0.0128	-0.0128
					(-2.97)		(-3.88)	(-3.26)	(-3.55)
φ_3						0.0698	0.0620	0.0334	
						(1.41)	(1.23)	(0.57)	
$\varphi 4$						0.0022	0.0024	0.0034	0.0022
						(2.39)	(2.51)	(2.15)	(2.42)
p_1	0.4756	0.4758	0.4754	0.4757	0.4758		0.4756	0.4757	0.4758
	(15.84)	(16.79)	(16.82)	(16.88)	(16.80)		(16.82)	(16.83)	(16.82)
c_1		0.2688		0.2685	0.2709		0.2709	0.2705	0.2718
		(14.32)		(14.31)	(14.44)		(14.45)	(14.42)	(14.41)
c_2		0.1785		0.1784	0.1790		0.1801	0.1787	0.1790
		(23.93)		(23.39)	(23.48)		(25.80)	(23.44)	(23.39)
c_3		-0.0031		-0.0037	-0.0043		-0.0032	-0.0041	
		(-0.44)		(-0.52)	(-0.45)		(-0.45)	(-0.46)	
c_4		0.0161		0.0158	0.0161		0.0281	0.0260	
		(1.25)		(1.76)	(1.86)		(1.86)	(1.84)	
c_5		0.2293		0.2299	0.2274		0.2047	0.2276	0.2271
		(6.62)		(5.97)	(6.57)		(6.47)	(6.57)	(6.53)
c_6		-0.4531		-0.4351	-0.4540		-0.4015	-0.4529	-0.4697
		(-13.09)		(-11.64)	(-13.12)		(-11.59)	(-13.09)	(-14.60)
σ_ε^2	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034	0.0034
	(4.51)	(4.80)	(4.88)	(4.83)	(4.87)	(4.85)	(4.86)	(4.94)	(4.90)
σ_ξ^2	0.00005	0.00005	0.00004	0.00005	0.00005	0.00005	0.00005	0.00005	0.00005
	(0.71)	(0.67)	(0.60)	(0.62)	(0.68)	(0.61)	(0.60)	(0.73)	(0.65)
J-stats	2.30	5.86	6.43	9.43	9.71	7.92	8.72	1.65	1.15
	(0.13)	(0.21)	(0.09)	(0.15)	(0.05)	(0.05)	(0.19)	(0.20)	(0.28)
$MSE_{in\ sample}$	0.7225	0.7214	0.7385	0.7216	0.7199	0.7231	0.7199	0.7195	0.7184
$MSE_{out\text{-}of\text{-}sample}$	1.1667	1.1657	1.1833	1.1653	1.1648	1.1671	1.1645	1.1604	1.1596

C. NP I

		Theta			Phi			Full	
	MRR	T.I	Risk	T.I+Risk	T.I	Risk	T.I+Risk	All	Significant
ϑ_1	0.1201	0.0285	0.0917	0.0314	0.1201	0.1201	0.1201	0.0195	
	(6.95)	(1.75)	(5.17)	(1.68)	(6.94)	(6.95)	(6.95)	(-1.47)	
$\vartheta1_{, uninformed}$		0.3450		0.3641				0.4634	0.3796
		(2.84)		(2.74)				(2.81)	(2.41)
$\vartheta1_{, fundamental}$		0.1729		0.1728				0.2347	0.2204
		(4.81)		(3.88)				(4.25)	(4.14)
$\vartheta1_{, informed}$		0.0630		0.0630				0.0735	0.0733
		(4.97)		(4.25)				(4.75)	(4.89)
$\vartheta2$			0.3072	-0.0165				0.0087	
			(1.87)	(-0.12)				(0.52)	
$\vartheta3$			0.0121	0.0105				-0.0107	
			(1.65)	(1.45)				(-0.89)	
φ_1	0.0369	0.0592	0.0763	0.0498	0.0491	0.0416	0.0539	0.1302	0.1103
	(2.01)	(3.64)	(4.53)	(3.02)	(2.69)	(2.70)	(2.67)	(4.32)	(4.06)
$\varphi_{2, uninformed}$					-0.0467		-0.0468	-0.3211	-0.2769
					(-2.26)		(-2.18)	(-2.14)	(-2.73)
$\varphi_{2, fundamental}$					-0.0320		-0.0192	-0.1684	-0.1538
					(-2.41)		(-2.34)	(-2.64)	(-2.57)
$\varphi_{2, informed}$					-0.0076		-0.0066	-0.0329	-0.0325
					(-2.79)		(-2.71)	(-2.65)	(-3.37)
φ_3						-0.1671	-0.1767	-0.1820	
						(-1.46)	(-1.64)	(-1.39)	
$\varphi4$						0.0127	0.0124	0.0192	0.0103
						(2.27)	(2.18)	(2.06)	(2.76)
p_1	0.2708	0.2660	0.2724	0.2709	0.2708	0.2709	0.2709	0.2710	0.2708
	(13.71)	(13.55)	(13.88)	(13.87)	(13.71)	(13.72)	(13.72)	(13.73)	(13.72)
c_1		0.7447		0.7464	0.8126		0.8123	0.8126	0.8919
		(3.94)		(3.95)	(4.41)		(4.43)	(4.41)	(8.25)
c_2		0.2457		0.2461	0.2582		0.2583	0.2588	0.2829
		(4.02)		(4.01)	(4.32)		(4.34)	(4.32)	(7.51)
c_3		-0.2988		-0.2984	-0.2960		-0.2967	-0.2897	-(0.30)
		(-2.51)		(-2.51)	(-2.42)		(-2.43)	(-2.45)	(-2.49)
c_4		-0.0443		-0.0393	-0.0173		-0.0172	-0.0168	
		(-0.52)		(-0.45)	(-0.20)		(-0.21)	(-0.19)	
c_5		0.1795		0.1545	0.1445		0.1444	0.1444	
		(0.55)		(0.47)	(0.46)		(0.46)	(0.65)	
c_6		-1.2096		-1.1872	-1.2677		-1.2679	-1.2576	-1.2723
		(-6.49)		(-6.46)	(-6.72)		(-6.69)	(-6.70)	(-6.72)
σ_ε^2	0.0046	0.0046	0.0039	0.0040	0.0046	0.0046	0.0046	0.0046	0.0046
	(2.03)	(1.98)	(1.97)	(1.80)	(2.13)	(2.06)	(2.16)	(2.01)	(2.01)
σ_ξ^2	0.00025	0.00024	0.00025	0.00023	0.00025	0.00025	0.00025	0.00024	0.00024
	(1.94)	(2.13)	(1.51)	(2.80)	(2.10)	(1.94)	(2.09)	(2.09)	(2.08)
J-stats	0.14	8.10	7.01	11.40	2.14	1.82	4.20	0.13	0.11
	(0.71)	(0.09)	(0.07)	(0.08)	(0.71)	(0.66)	(0.65)	(0.72)	(0.74)
$MSE_{in\ sample}$	0.6907	0.6394	0.6372	0.6346	0.6380	0.6384	0.6362	0.6247	0.6140
$MSE_{out\text{-}of\text{-}sample}$	1.5114	1.3598	3.4869	1.3362	1.5865	1.6670	1.4127	0.9561	0.9494

D. NP II

		Theta			Phi			Full	
	MRR	T.I	Risk	T.I+Risk	T.I	Risk	T.I+Risk	All	Significant
ϑ_1	0.0657	0.0508	0.1385	0.0650	0.0666	0.0657	0.0665	-0.0655	
	(2.38)	(1.56)	(3.47)	(1.57)	(2.44)	(2.38)	(2.43)	(-0.95)	
$\vartheta_{2,\ uninformed}$		0.0421		0.0412				0.7702	0.7121
		(2.22)		(1.21)				(2.92)	(3.02)
$\vartheta_{2,\ fundamental}$		0.0046		0.0034				0.6868	0.5591
		(2.01)		(2.32)				(3.05)	(2.99)
$\vartheta_{2,\ informed}$		0.0015		0.0013				0.0142	0.0295
		(2.28)		(2.62)				(3.20)	(3.59)
ϑ_3			-0.4164	-0.3575				0.2454	
			(-1.88)	(-1.62)				(0.61)	
ϑ_4			0.0000	0.0000				0.0000	
			(-1.05)	(-0.92)				(-0.98)	
φ_1	0.0597	0.0656	0.0371	0.0464	0.0700	0.1272	0.1074	0.1227	0.1080
	(2.09)	(2.45)	(1.97)	(2.01)	(2.31)	(2.81)	(3.12)	(3.59)	(4.97)
$\varphi_{2,\ uninformed}$					-0.0716		-0.0701	-0.0961	-0.0729
					(-2.05)		(-1.95)	(-2.56)	(-2.56)
$\varphi_{2,\ fundamental}$					-0.0220		-0.0229	-0.0464	-0.0445
					(-2.28)		(-2.29)	(-2.96)	(-2.89)
$\varphi_{2,\ informed}$					-0.0079		-0.0074	-0.0177	-0.0169
					(-3.33)		(-3.37)	(-3.03)	(-2.98)
φ_3						-0.5408	-0.5404	-0.6984	-0.3585
						(-2.68)	(-2.77)	(-2.03)	(2.80)
$\varphi 4$						0.0000	0.0000	0.0000	
						(-0.42)	(-0.61)	(0.46)	
p_1	0.2946	0.2965	0.2950	0.2973	0.2945	0.2946	0.2945	0.2945	0.2944
	(14.68)	(14.80)	(14.70)	(14.84)	(14.67)	(14.68)	(14.64)	(14.60)	(14.66)
c_1		0.4027		0.3974	0.4584		0.4583	0.4581	0.4717
		(2.69)		(2.66)	(3.13)		(3.15)	(3.12)	(3.35)
c_2		0.1530		0.1508	0.1700		0.1690	0.1696	0.1661
		(2.62)		(2.59)	(2.92)		(2.94)	(2.91)	(3.10)
c_3		0.0867		0.0877	0.0863		0.0856	0.0862	
		(1.41)		(1.43)	(1.49)		(1.43)	(1.43)	
c_4		-0.0410		-0.0375	-0.0576		-0.0575	-0.0570	
		(-0.53)		(-0.49)	(-0.78)		(-0.83)	(-0.77)	
c_5		0.6094		0.6128	0.5873		0.5838	0.5896	0.5765
		(2.73)		(2.74)	(2.71)		(2.68)	(2.70)	(2.76)
c_6		-1.1298		-1.1261	-1.0856		-1.1084	-1.0856	-1.1223
		(6.17)		(-6.18)	(-6.12)		(-6.37)	(-6.11)	(-6.41)
σ_{ε}^2	0.0018	0.0019	0.0020	0.0019	0.0018	0.0019	0.0019	0.0019	0.0019
	(2.54)	(2.44)	(2.08)	(2.86)	(2.52)	(2.47)	(2.51)	(2.42)	(2.40)
σ_{ξ}^2	0.00016	0.00015	0.00014	0.00019	0.00018	0.00015	0.00015	0.00015	0.00015
	(1.73)	(1.99)	(1.79)	(1.96)	(1.73)	(1.81)	(1.81)	(1.90)	(1.89)
J-stats	0.24	5.43	4.67	10.66	2.54	1.75	4.32	0.08	0.04
	(0.63)	(0.25)	(0.20)	(0.10)	(0.63)	(0.63)	(0.63)	(0.78)	(0.85)
$MSE_{in\ sample}$	0.4761	0.3829	0.3842	0.3800	0.4497	0.4690	0.4438	0.3798	0.3777
$MSE_{out\text{-}of\text{-}sample}$	0.6378	0.6019	0.6135	0.6001	0.6247	0.6341	0.6192	0.5995	0.5873

Table 6.20: Information and Liquidity Spread Components

		Average Trading Intensity	Theta	Average Theta	Phi	Average Phi	Implied Spread (No Risk)
ECX I					0.0969		
	Uninformed	0.5405	0.0987	0.0533	-0.0889	0.0489	0.2951
	Fundamental	1.6873	0.0469	0.0791	-0.0539	0.0060	0.1058
	Informed	2.4551	0.0398	0.0977	-0.0127	0.0657	0.2110
ECX II					0.0599		
	Uninformed	0.6945	0.0613	0.0426	-0.0522	0.0237	0.1699
	Fundamental	1.3916	0.0391	0.0544	-0.0385	0.0063	0.0908
	Informed	2.7120	0.0286	0.0776	-0.0128	0.0252	0.1076
NP I					0.1103		
	Uninformed	0.5405	0.3796	0.2052	-0.2769	-0.0394	0.6805
	Fundamental	1.6873	0.2204	0.3719	-0.1538	-0.1492	0.1424
	Informed	2.4551	0.0733	0.1800	-0.0325	0.0305	0.2076
NP II					0.108		
	Uninformed	0.5405	0.7121	0.3849	-0.0729	0.0686	1.5614
	Fundamental	1.6873	0.5591	0.9433	-0.0445	0.0329	1.1840
	Informed	2.4551	0.0295	0.0724	-0.0169	0.0665	0.1920

Table 6.21: Intraday Variations of Spread, Variance and their Components

A. ECX

ECX I

Time of the Day	Trading Intensity	Expected Trading Intensity	Adverse selection	Liquidity	Risk Aversion	Phi	Implied Spread	Effective Spread	Assymetric Information	Trading Costs	Interaction	Return Variance
07	1.2700	1.1700	0.0752	0.0492	0.0043	0.0534	0.2573	0.1821	0.0061	0.0028	0.0032	0.0206
08	1.2600	1.1674	0.0721	0.0433	0.0043	0.0476	0.2394	0.1673	0.0052	0.0028	0.0036	0.0202
09	0.9883	1.0521	0.0662	0.0468	0.0047	0.0515	0.2354	0.1692	0.0044	0.0032	0.0035	0.0197
10	0.9668	0.9517	0.0606	0.0506	0.0043	0.0549	0.2310	0.1704	0.0040	0.0036	0.0033	0.0194
11	0.9656	0.8681	0.0554	0.0547	0.0053	0.0600	0.2307	0.1753	0.0035	0.0039	0.0033	0.0193
12	1.0800	1.0108	0.0621	0.0509	0.0056	0.0565	0.2371	0.1750	0.0042	0.0035	0.0035	0.0199
13	1.0828	1.0179	0.0634	0.0490	0.0048	0.0537	0.2342	0.1709	0.0043	0.0034	0.0034	0.0197
14	0.8991	0.9359	0.0600	0.0505	0.0057	0.0561	0.2322	0.1723	0.0038	0.0035	0.0033	0.0193
15	1.0440	0.8972	0.0563	0.0539	0.0059	0.0598	0.2323	0.1759	0.0035	0.0039	0.0033	0.0193
16	1.1500	1.0500	0.0520	0.0662	0.0093	0.0755	0.2550	0.2030	0.0037	0.0045	0.0036	0.0204

ECX II

Time of the Day	Trading Intensity	Expected Trading Intensity	Adverse selection	Liquidity	Risk Aversion	Phi	Implied Spread	Effective Spread	Assymetric Information	Trading Costs	Interaction	Return Variance
07	1.2400	1.1600	0.0570	0.0144	0.0070	0.0174	0.1487	0.0917	0.0023	0.0004	0.0009	0.0071
08	1.0205	1.0490	0.0495	0.0148	0.0030	0.0164	0.1317	0.0822	0.0020	0.0005	0.0008	0.0067
09	1.0489	1.0137	0.0480	0.0160	0.0016	0.0188	0.1336	0.0856	0.0019	0.0005	0.0008	0.0066
10	0.9156	0.9685	0.0464	0.0176	0.0028	0.0209	0.1346	0.0881	0.0018	0.0005	0.0009	0.0066
11	0.9665	0.9170	0.0467	0.0175	0.0032	0.0226	0.1386	0.0919	0.0019	0.0006	0.0009	0.0068
12	0.8500	0.8900	0.0475	0.0166	0.0052	0.0217	0.1384	0.0909	0.0019	0.0006	0.0009	0.0068
13	1.0002	1.0161	0.0481	0.0159	0.0036	0.0195	0.1352	0.0871	0.0020	0.0005	0.0009	0.0068
14	1.0767	1.0274	0.0484	0.0155	0.0030	0.0186	0.1340	0.0856	0.0019	0.0005	0.0008	0.0067
15	1.0205	0.9758	0.0459	0.0172	0.0038	0.0210	0.1339	0.0880	0.0016	0.0006	0.0010	0.0065
16	1.4000	1.1900	0.0501	0.0257	0.0068	0.0325	0.1653	0.1152	0.0019	0.0009	0.0015	0.0077

Appendix 6.B

B. NP

NP I

Time of the Day	Trading Intensity	Expected Trading Intensity	Adverse selection	Liquidity	Risk Aversion	Phi	Implied Spread	Effective Spread	Assymetric Information	Trading Costs	Interaction	Return Variance
08	1.4795	1.1502	0.1560	0.0359	0.0022	0.0380	0.3880	0.2320	0.0226	0.0021	0.0110	0.0408
09	1.4503	1.1431	0.1417	0.0257	0.0070	0.0327	0.3488	0.2071	0.0186	0.0016	0.0086	0.0339
10	1.2663	0.9997	0.1294	0.0318	0.0069	0.0387	0.3363	0.2069	0.0155	0.0022	0.0093	0.0321
11	0.9890	0.9360	0.1318	0.0282	0.0095	0.0378	0.3392	0.2074	0.0161	0.0021	0.0092	0.0325
12	0.8395	0.7098	0.1150	0.0384	0.0140	0.0524	0.3348	0.2198	0.0123	0.0032	0.0091	0.0296
13	1.1651	0.9900	0.1302	0.0312	0.0131	0.0443	0.3491	0.2189	0.0157	0.0029	0.0107	0.0344
14	0.9272	1.0058	0.1304	0.0400	0.0134	0.0534	0.3676	0.2372	0.0158	0.0042	0.0129	0.0379
15	1.0927	1.1258	0.1375	0.0401	0.0138	0.0539	0.3828	0.2453	0.0175	0.0042	0.0137	0.0406

NP II

Time of the Day	Trading Intensity	Expected Trading Intensity	Adverse selection	Liquidity	Risk Aversion	Phi	Implied Spread	Effective Spread	Assymetric Information	Trading Costs	Interaction	Variance
08	1.1072	1.1640	0.1140	0.0524	0.0089	0.0613	0.3505	0.2365	0.0148	0.0066	0.0133	0.0370
09	0.9627	0.9972	0.0714	0.0553	0.0226	0.0779	0.2987	0.2273	0.0058	0.0107	0.0127	0.0300
10	1.0119	0.7915	0.0649	0.0577	0.0216	0.0794	0.2885	0.2236	0.0048	0.0112	0.0118	0.0277
11	0.9801	0.9725	0.0697	0.0537	0.0273	0.0810	0.3013	0.2316	0.0055	0.0116	0.0129	0.0300
12	1.0418	0.9580	0.0682	0.0550	0.0261	0.0810	0.2985	0.2303	0.0053	0.0116	0.0126	0.0310
13	1.0125	1.0884	0.0751	0.0504	0.0398	0.0902	0.3305	0.2555	0.0064	0.0144	0.0155	0.0363
14	0.9550	0.9035	0.0603	0.0591	0.0352	0.0943	0.3092	0.2488	0.0042	0.0157	0.0130	0.0310
15	1.1209	1.1271	0.0814	0.0637	0.0324	0.0961	0.3548	0.2735	0.0076	0.0163	0.0150	0.0420

Table 6.22: Spread and Variance across different regimes of Transaction size, Trading Frequency and Trade Initiation

A. Spread

		ECX I				ECX II			
		Previous Trade		Contemporaneous Trade		Previous Trade		Contemporaneous Trade	
		Adverse Selection	Implied Spread	Adverse Selection	Implied Spread	Adverse Selection	Implied Spread	Adverse Selection	Implied Spread
B	0-50	0.0491	0.2331	0.0731	0.2332	0.0496	0.1307	0.0497	0.1301
	51-100	0.0569	0.2445	0.0725	0.2444	0.0532	0.1356	0.0516	0.1329
	>100	0.0735	0.2473	0.1015	0.2801	0.0574	0.1338	0.0527	0.1335
O	0-50	0.0331	0.2189	0.0556	0.2195	0.0468	0.1305	0.0470	0.1305
	51-100	0.0430	0.2286	0.0473	0.2175	0.0490	0.1334	0.0442	0.1317
	>100	0.0558	0.2297	0.0457	0.2120	0.0536	0.1354	0.0448	0.1323
B	0-500	0.0783	0.2373	0.0759	0.2359	0.0501	0.1310	0.0498	0.1385
	501-1000	0.0366	0.2022	0.0548	0.2154	0.0296	0.1295	0.0419	0.1347
	>1000	0.0251	0.2017	0.0480	0.2116	0.0259	0.1283	0.0408	0.1377
O	0-500	0.0613	0.2228	0.0574	0.2209	0.0474	0.1306	0.0471	0.1351
	501-1000	0.0290	0.2019	0.0480	0.2125	0.0283	0.1294	0.0407	0.1340
	>1000	0.0190	0.2023	0.0380	0.2105	0.0249	0.1281	0.0401	0.1309

		NP I				NP II			
B	0-50	0.1284	0.3404	0.1309	0.3428	0.1218	0.4923	0.1307	0.5093
	51-100	0.1369	0.3564	0.1408	0.3456	0.4437	1.0831	0.2219	0.6311
	>100	0.1565	0.4387	0.1980	0.4419	1.1244	2.3842	0.2303	0.6383
O	0-50	0.1180	0.3316	0.1129	0.3008	0.0624	0.3828	0.1187	0.4782
	51-100	0.1309	0.3425	0.1155	0.3274	0.0773	0.4039	0.1451	0.5106
	>100	0.1446	0.3693	0.1282	0.3403	0.1208	0.4816	0.1988	0.5699
B	0-500	0.1494	0.3894	0.1449	0.3549	0.1642	0.5444	0.2020	0.5407
	501-1000	0.1252	0.3180	0.1290	0.3267	0.0987	0.4537	0.0624	0.5209
	>1000	0.1024	0.3160	0.1166	0.3058	0.0946	0.4810	0.0595	0.3726
O	0-500	0.1434	0.3582	0.1400	0.3540	0.1323	0.4834	0.1773	0.5007
	501-1000	0.1166	0.3117	0.1223	0.3241	0.1182	0.5030	0.0603	0.3704
	>1000	0.1031	0.3097	0.1184	0.3057	0.1019	0.4662	0.0486	0.3012

B. Variance

		ECX I				ECX II			
		Previous Trade		Contemporaneous Trade		Previous Trade		Contemporaneous Trade	
		Assymetric Information	Return Variance	Assymetric Information	Return Variance	Assymetric Information	Return Variance	Assymetric Information	Return Variance
B	0-50	0.0033	0.0190	0.0052	0.0184	0.0021	0.0069	0.0021	0.0069
	51-100	0.0045	0.0208	0.0056	0.0192	0.0024	0.0070	0.0023	0.0069
	>100	0.0052	0.0210	0.0085	0.0195	0.0028	0.0077	0.0022	0.0070
O	0-50	0.0011	0.0203	0.0026	0.0203	0.0019	0.0068	0.0015	0.0068
	51-100	0.0033	0.0222	0.0031	0.0217	0.0020	0.0072	0.0016	0.0070
	>100	0.0034	0.0224	0.0034	0.0249	0.0024	0.0077	0.0019	0.0070
B	0-500	0.0057	0.0193	0.0055	0.0192	0.0021	0.0071	0.0023	0.0074
	501-1000	0.0012	0.0176	0.0032	0.0184	0.0007	0.0070	0.0017	0.0073
	>1000	0.0006	0.0180	0.0026	0.0184	0.0006	0.0070	0.0014	0.0072
O	0-500	0.0040	0.0207	0.0037	0.0206	0.0019	0.0073	0.0021	0.0076
	501-1000	0.0008	0.0174	0.0026	0.0183	0.0007	0.0070	0.0015	0.0072
	>1000	0.0004	0.0177	0.0018	0.0186	0.0005	0.0071	0.0014	0.0071

		NP I				NP II			
B	0-50	0.0213	0.0329	0.0213	0.0272	0.0176	0.0488	0.0655	0.0836
	51-100	0.0236	0.0356	0.0231	0.0317	0.0245	0.0526	0.0659	0.0999
	>100	0.0272	0.0381	0.0363	0.0366	0.0273	0.1132	0.0736	0.2713
O	0-50	0.0139	0.0373	0.0131	0.0357	0.0055	0.1199	0.0350	0.1120
	51-100	0.0189	0.0399	0.0152	0.0376	0.0099	0.1442	0.0528	0.1124
	>100	0.0208	0.0518	0.0189	0.0506	0.0107	0.1518	0.0582	0.1204
B	0-500	0.0290	0.0386	0.0243	0.0518	0.0852	0.1337	0.0846	0.1318
	501-1000	0.0205	0.0335	0.0227	0.0399	0.0829	0.1175	0.0753	0.0957
	>1000	0.0136	0.0303	0.0180	0.0381	0.0318	0.0754	0.0409	0.0739
O	0-500	0.0242	0.0419	0.0239	0.0526	0.0431	0.1536	0.0487	0.1640
	501-1000	0.0167	0.0355	0.0213	0.0399	0.0423	0.0941	0.0451	0.1075
	>1000	0.0158	0.0292	0.0136	0.0373	0.0109	0.0861	0.0294	0.0746

Appendix 6.C

Figures

Empirical Findings

The figures presented in the following section are based on basic statistics and examine various model implications related to price changes, spreads, price change variance and their components. There are 8 figures per market and they are organized across markets and phases. This results in 32 figures in total, where Figures 6.1-6.8 are related to ECX I, 6.9-6.16 to ECX II, 6.17-6.24 to NP I and 6.25-6.32 to NP II, where ECX I, ECX II, NP I and NP II stand for European Climate Exchange Phase I, European Climate Exchange Phase II, Nord Pool Phase I and Nord Pool Phase II, respectively.

Focusing on each market, the first group of figures (6.1, 6.9, 6.17 and 6.25) presents the intraday variations (hours are expressed in 24h format) of Trading Intensity, s_t, Expected Trading Intensity, $E[s_t|H_{t-1}]$, Implied Spread, $S_t^{Implied}$, and Effective Spread, $S_t^{Effective}$, (Figure A), of the Components of the Implied spread (Figure B) and of Price Change Variance, $Var(\Delta p_t)$, and its Components (Figure C).

Along the same lines, the second group of figures (6.2, 6.10, 6.18 and 6.26) focuses on the intraday formation of Implied Spread, Price Change Variance and their components in absolute numbers (Figures A and C) and as a proportion of the total values (Figures B and D).

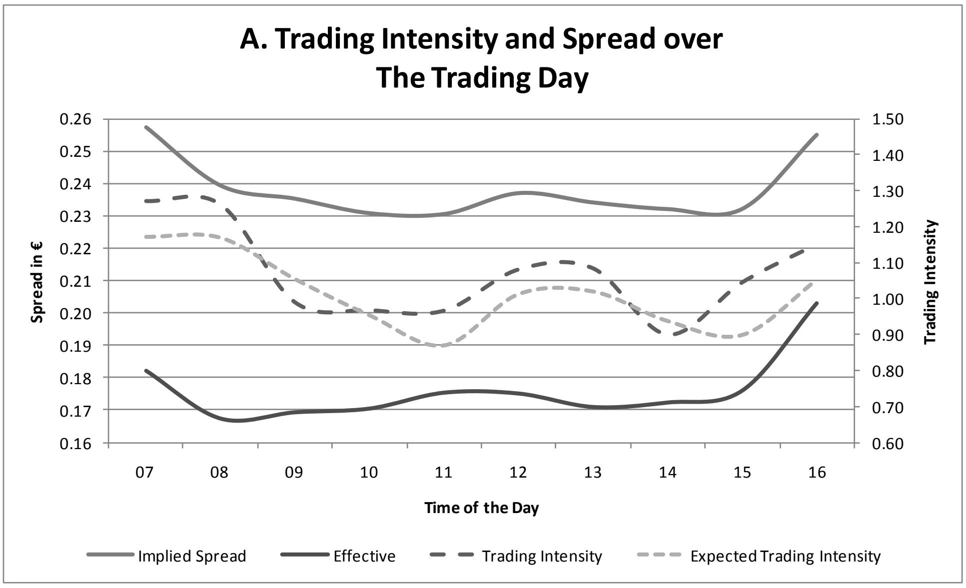
The third group of figures (6.3, 6.11, 6.19 and 6.27) presents the average Implied spread (Figure A) and Price Change Variance (Figure B) across Expected Trading intensity (the values are calculated according to Eq. (6.22)) and during the trading day.

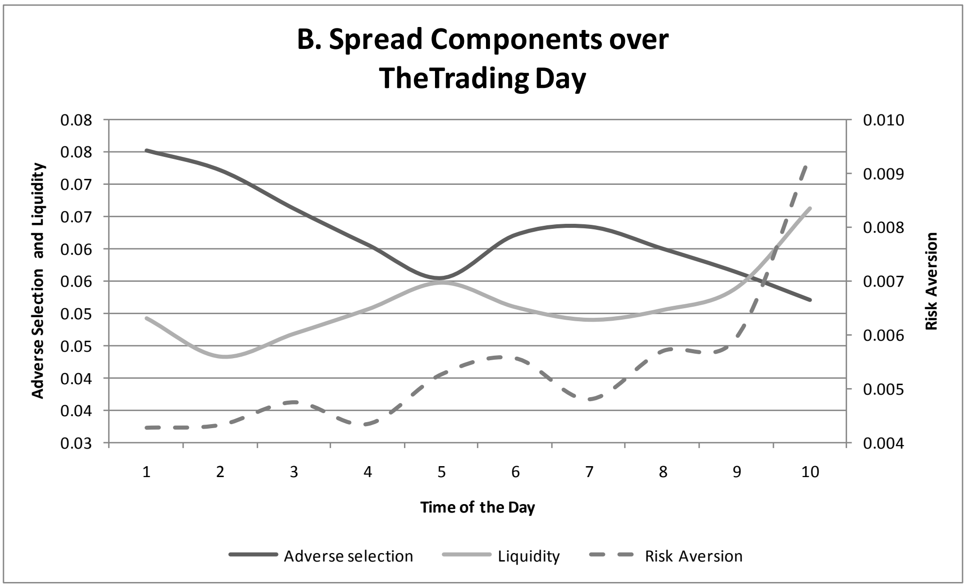
The fourth (6.4, 6.12, 6.20 and 6.28) and fifth (6.5, 6.13, 6.21 and 6.29) group of figures present the variations of Implied Spread and Price Change Variance, respectively, across expectations. More specifically, Figure A examines the variations across Expected Trading Intensity and Expected Informed Trading, noted as “Informed Trading”, $E[\tilde{I}_t|H_{t-1}]$, formulated as in Eq. (6.14). Figure B emphasizes the relative importance of the expected presence of OTC traders, noted as “OTC Transactions”, which is measured by their proportion of total trading in the last 15 minutes, along with the Expected Trading Intensity. Figure C examines the spread’s and variance’s variations across Expected Trading Intensity and Expected Price Volatility, noted as “Price Volatility”, $E\left[\sigma_{p,t}^2|H_{t-1}\right]$, formulated as in Eq. (6.13)

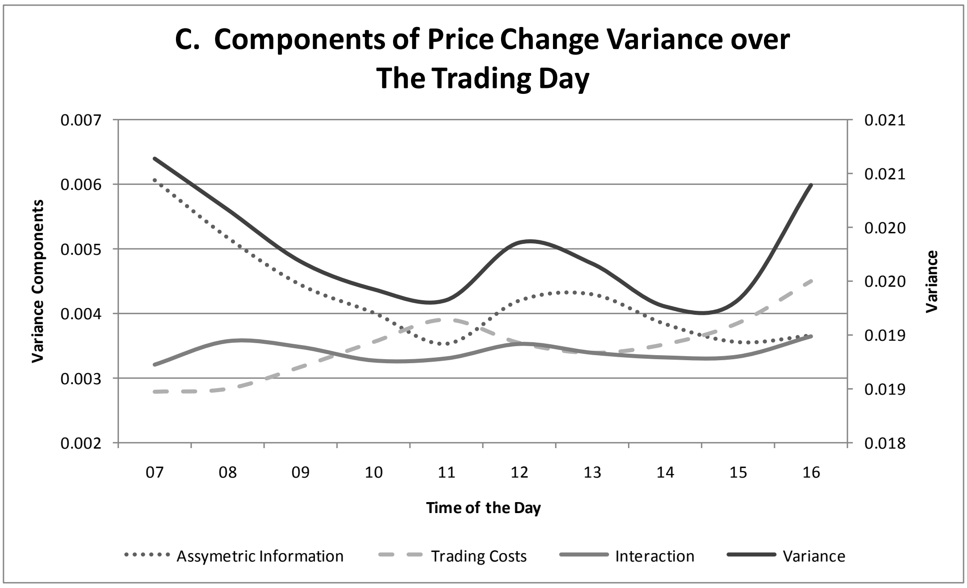
The sixth (6.6, 6.14, 6.22 and 6.30) and seventh (6.7, 6.15, 6.23 and 6.31) group of figures present the variations of the Adverse-selection, θ_t, and Liquidity, φ_t, component of the Implied Spread, respectively, across expectations. Figure A focuses on the Expected Trading Intensity and the Expected Informed Trading, Figure B focuses on the Expected Trading Intensity and the expected presence of OTC traders, while Figure C focuses on the Expected Trading Intensity and the Expected Price Volatility.

Finally, the last group of figures (6.8, 6.16, 6.24 and 6.32) presents the difference between the rate of change of the Implied Spread and the rate of change of the Price Change Variance, R_t, computed according to Eq. (6.31), across Expected Trading Intensity and Expected Informed Trading (Figure A), Expected presence of OTC traders (Figure B) and Expected Price Volatility (Figure C).

Figure 6.6: ECX I Intraday Variations of Spread, Variance and Their Components







Appendix 6.C

Figure 6.7: ECX I Intraday Spread and Variance Components

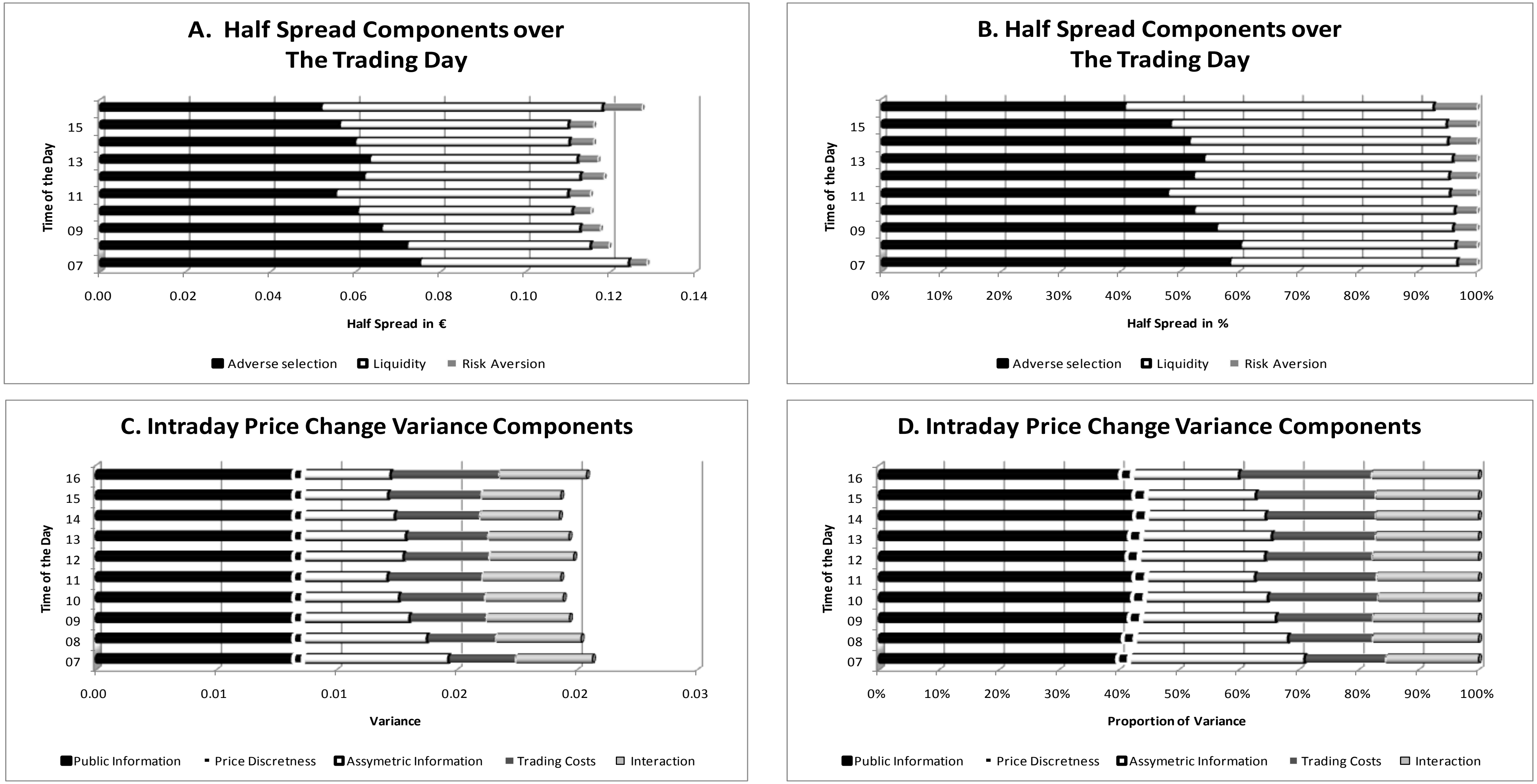
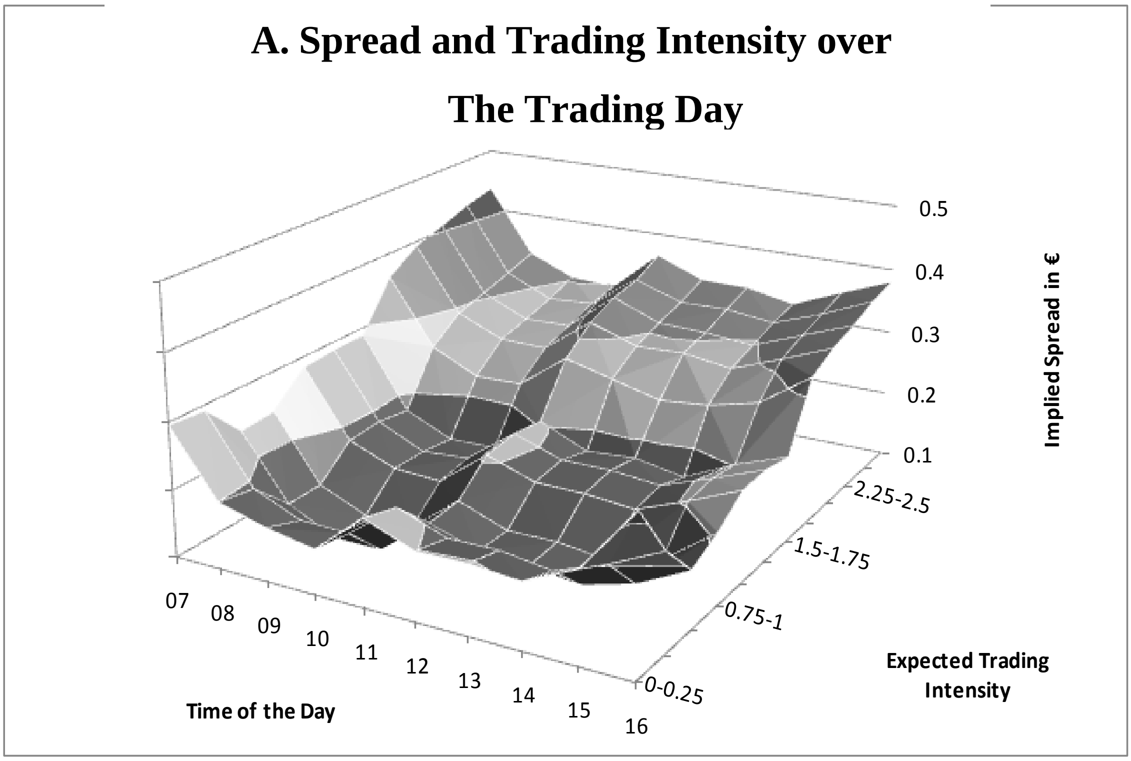


Figure 6.8: ECX I Spread and Variance over the Trading Day and across Trading Intensity



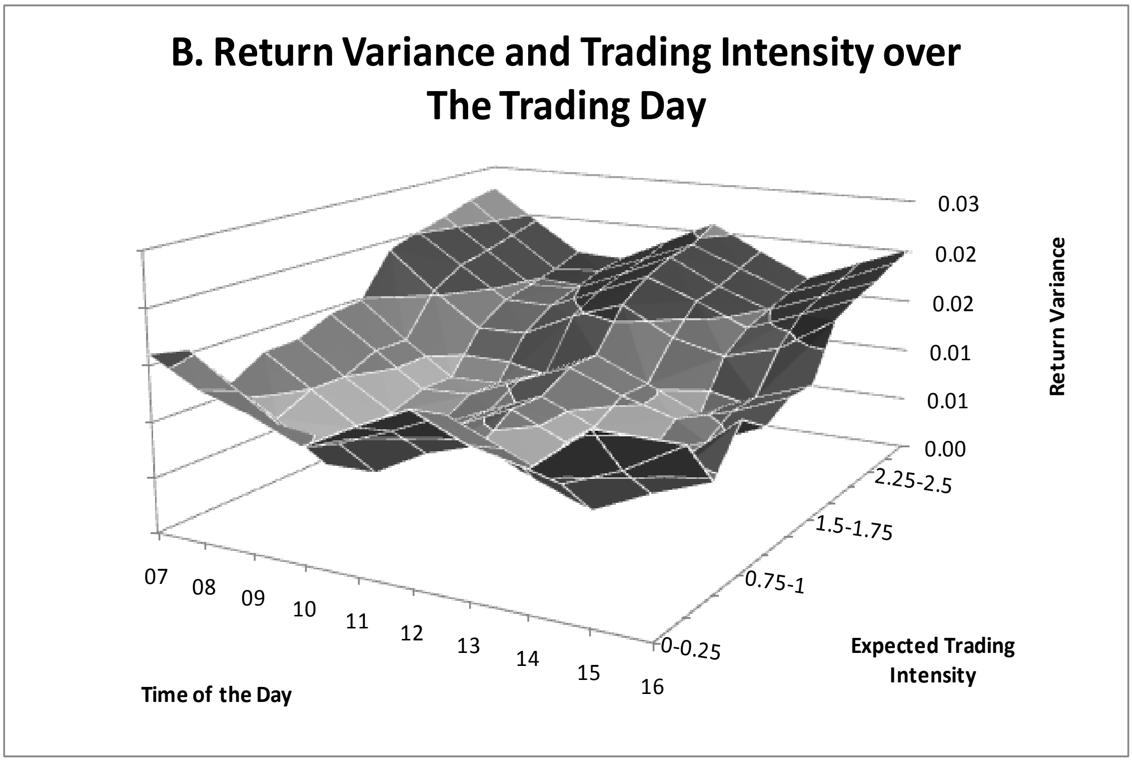
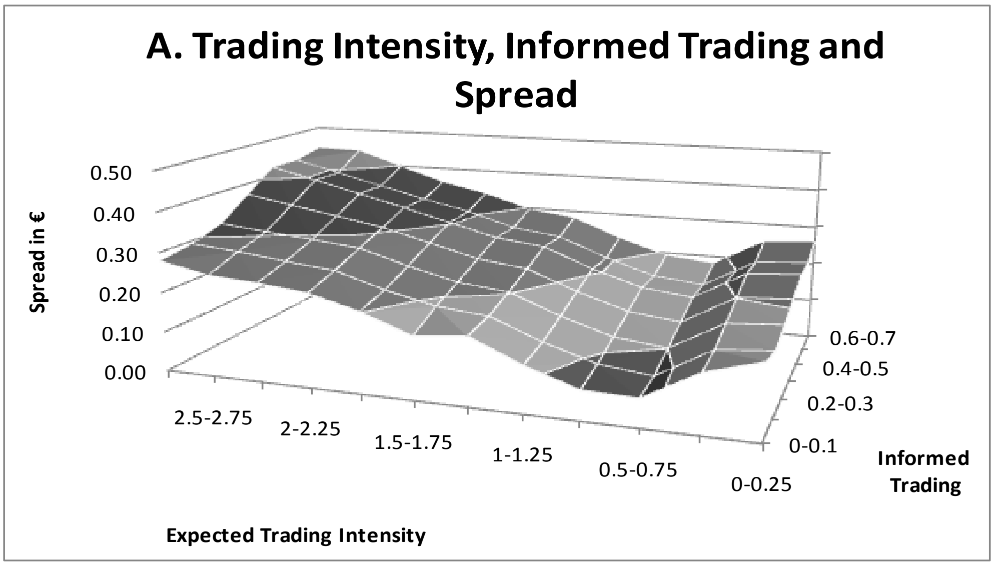
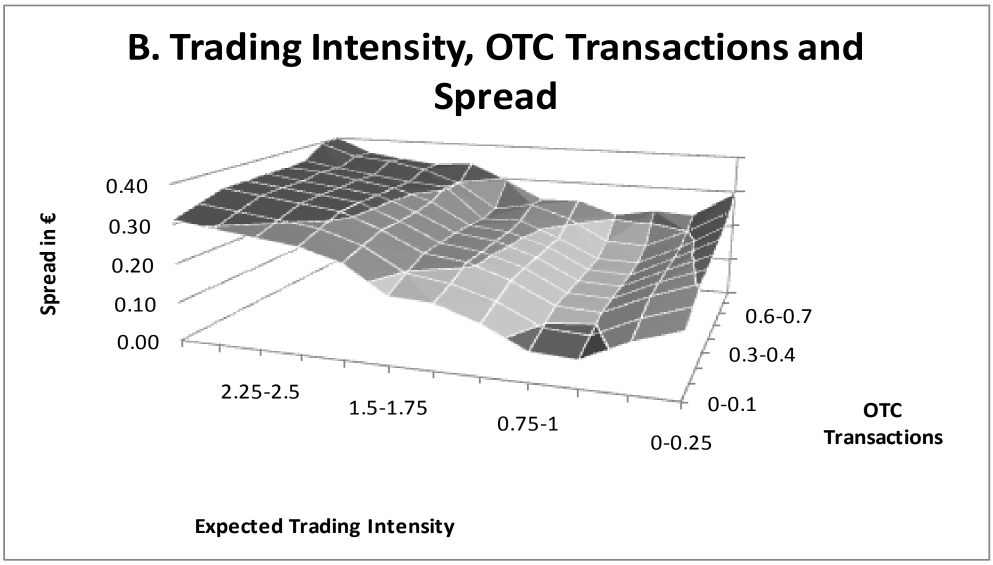


Figure 6.4: ECX I Spread across Trading Intensity, Risk and OTC Transactions





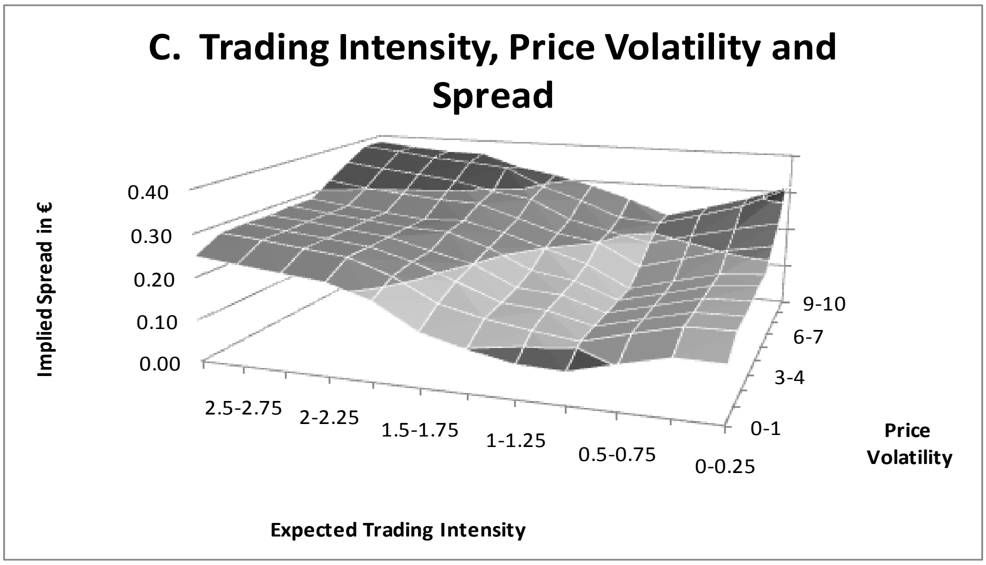
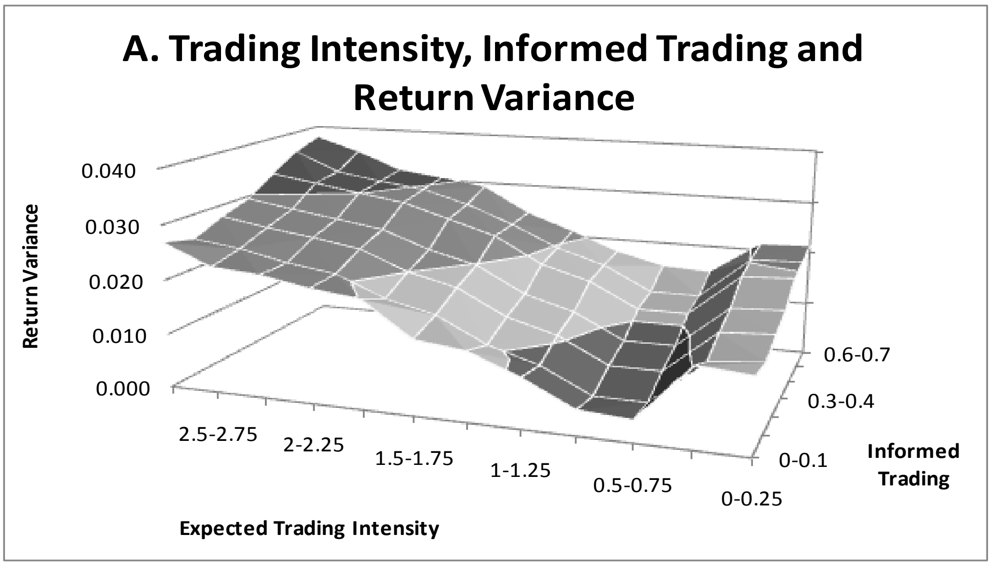
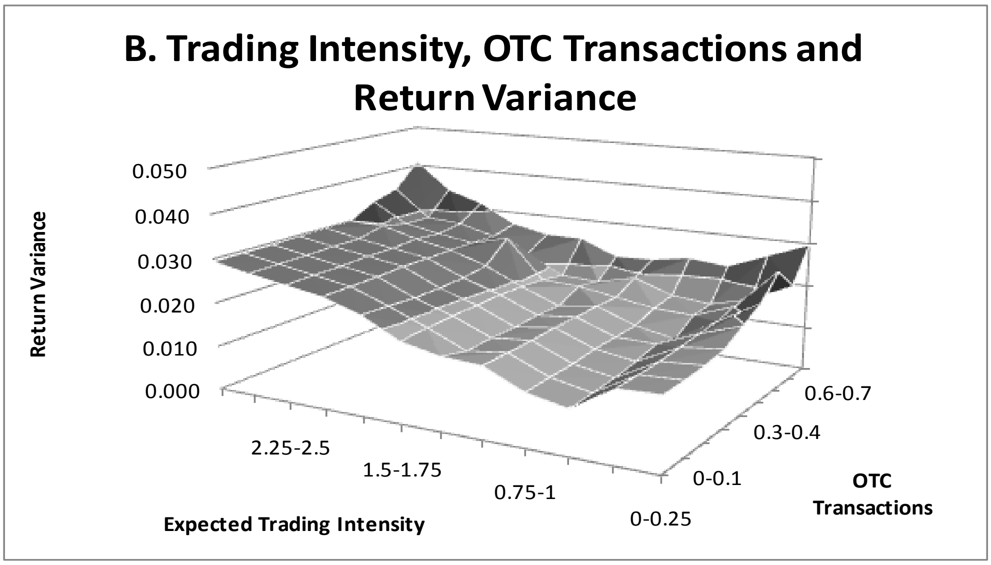


Figure 6.5: ECX I Variance across Trading Intensity, Risk and OTC Transactions





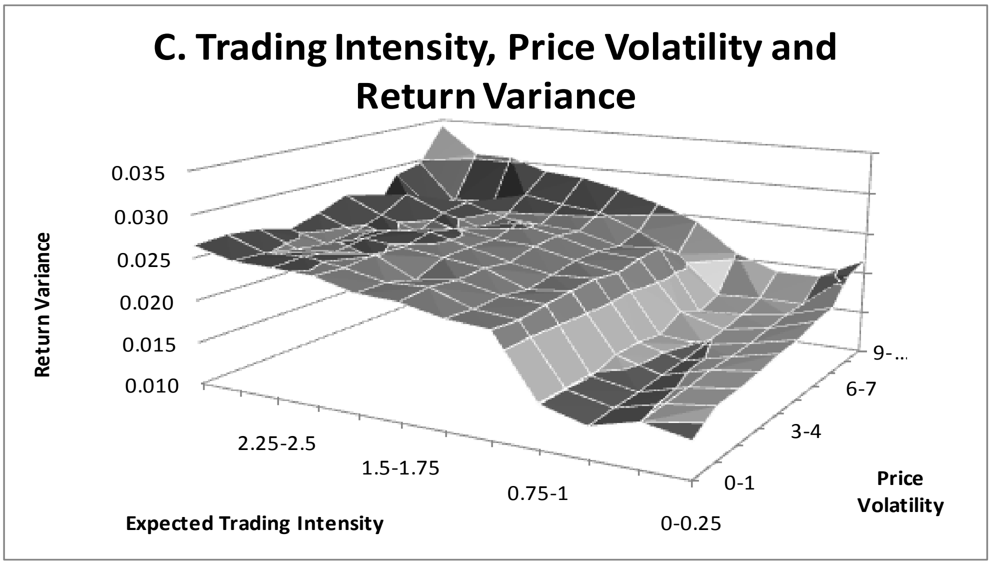
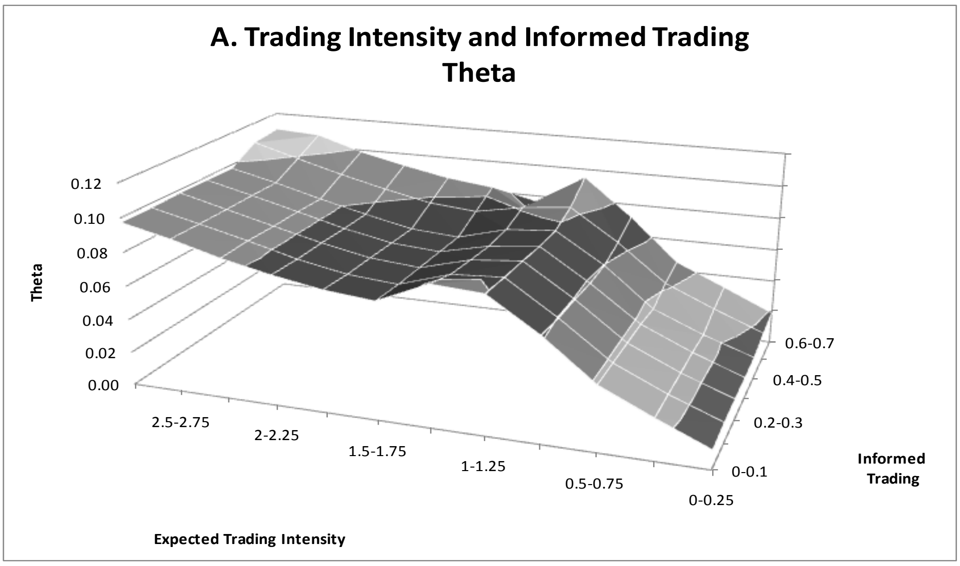
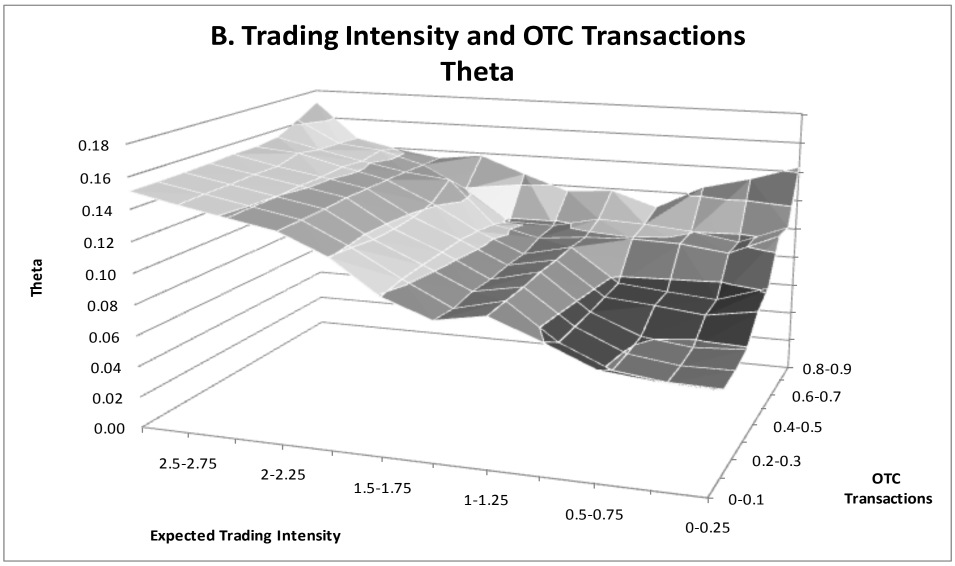


Figure 6.6: ECX I Theta across Trading Intensity, Risk and OTC Transactions





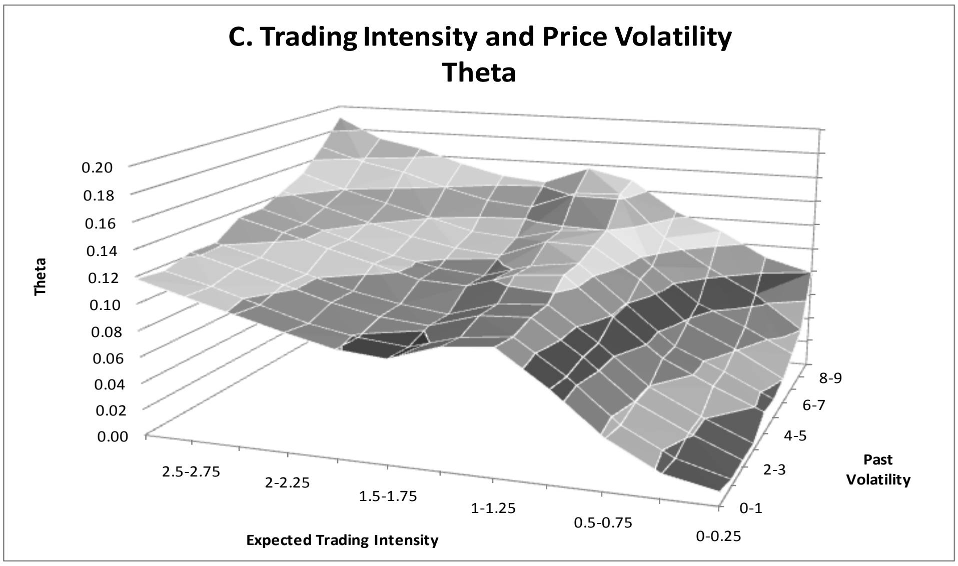
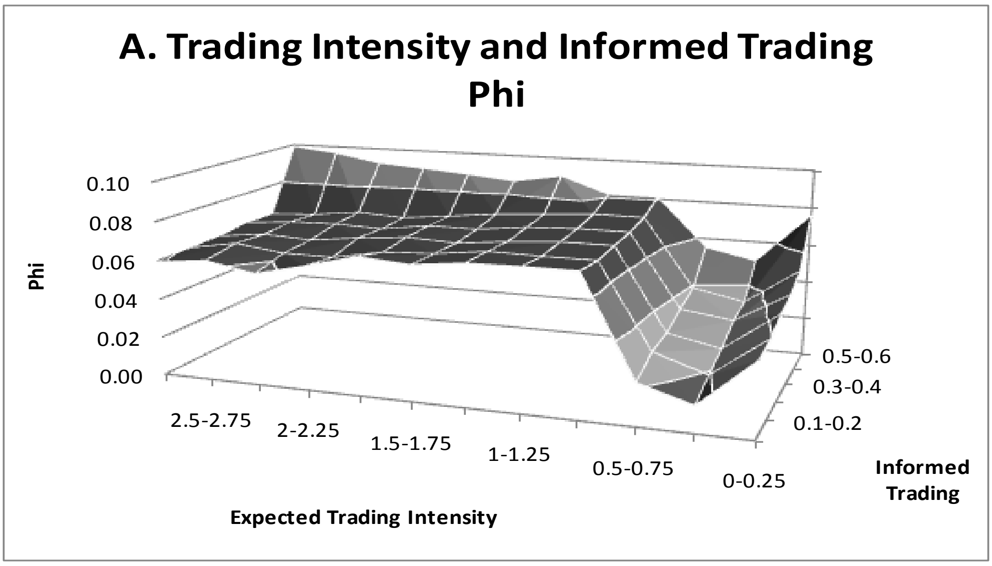
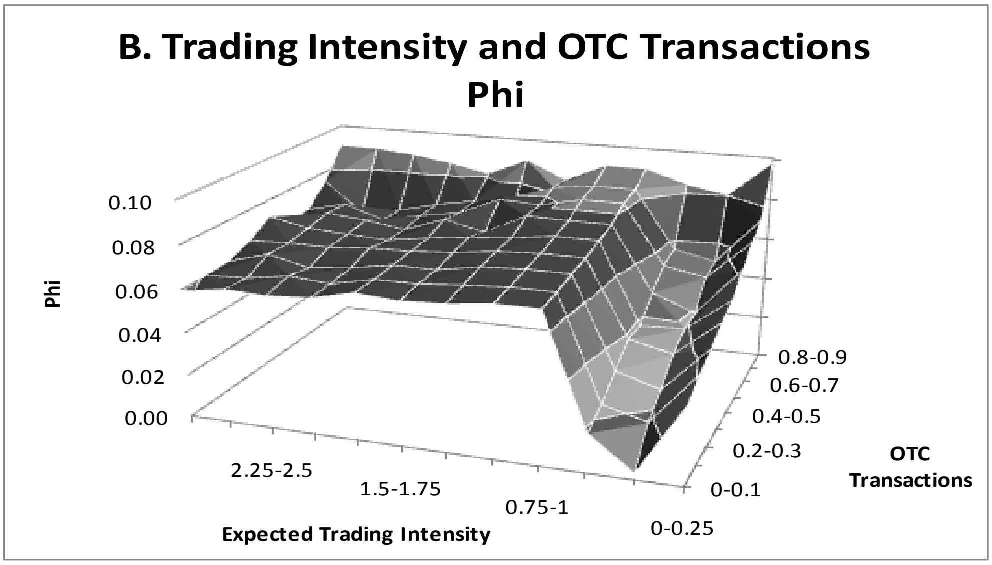


Figure 6.7: ECX I Phi across Trading Intensity, Risk and OTC Transactions





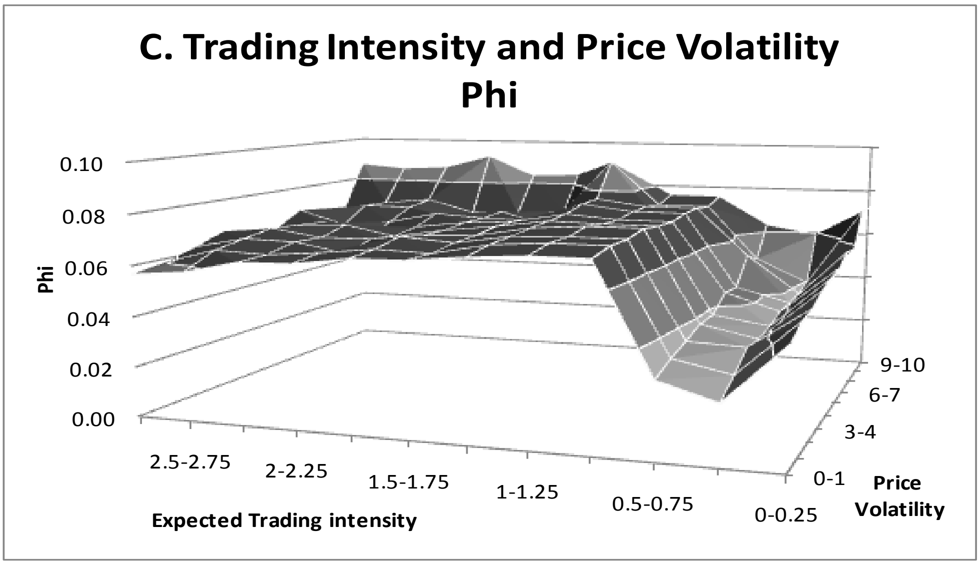
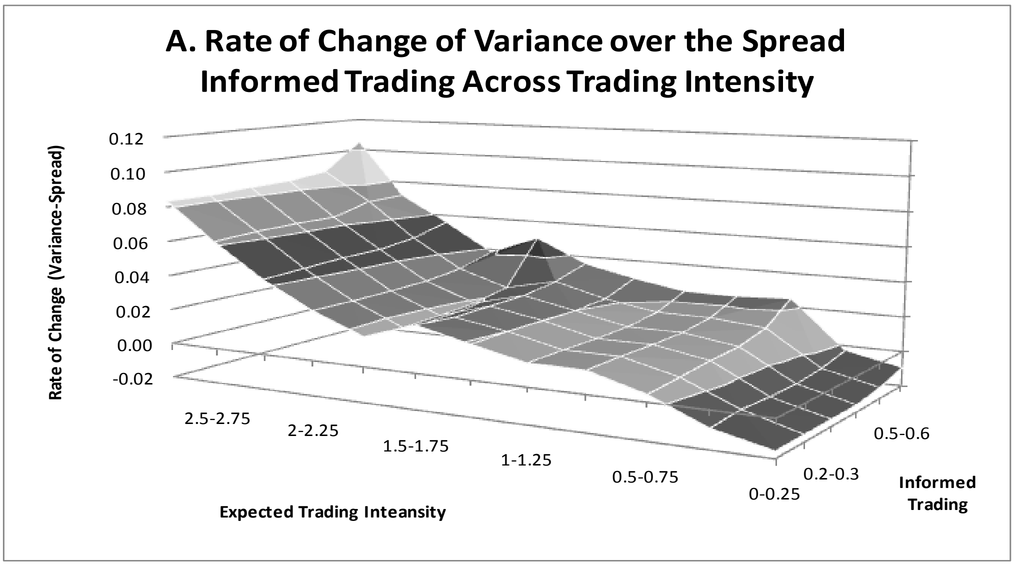
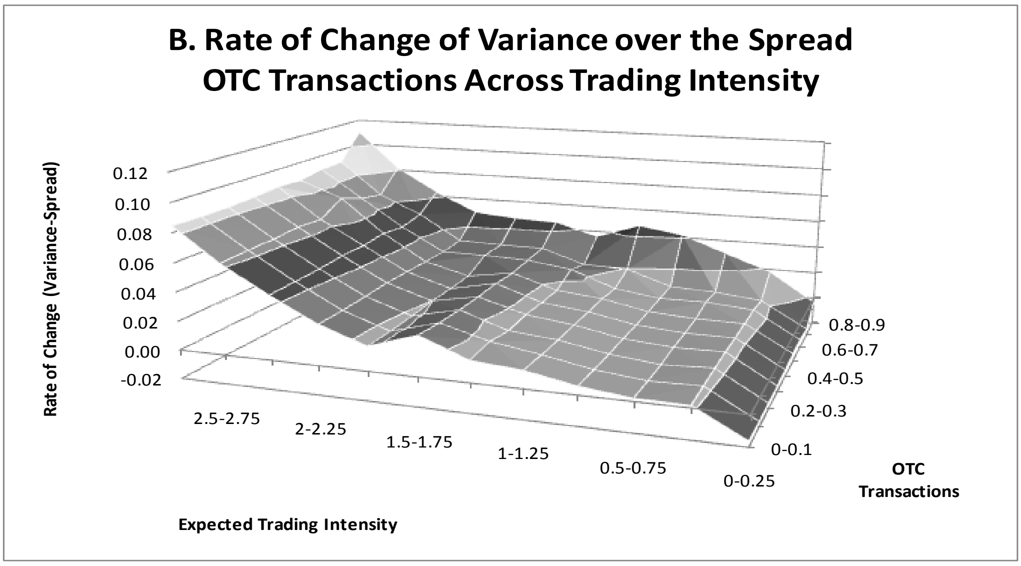


Figure 6.8: ECX I Marginal Variance over Spread change across Trading Intensity, Risk and OTC Transactions





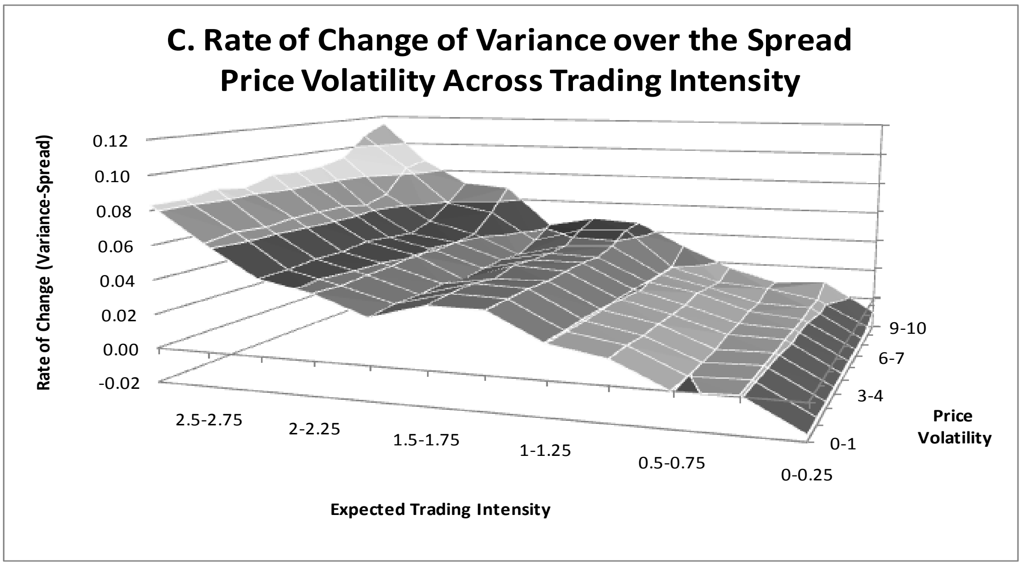
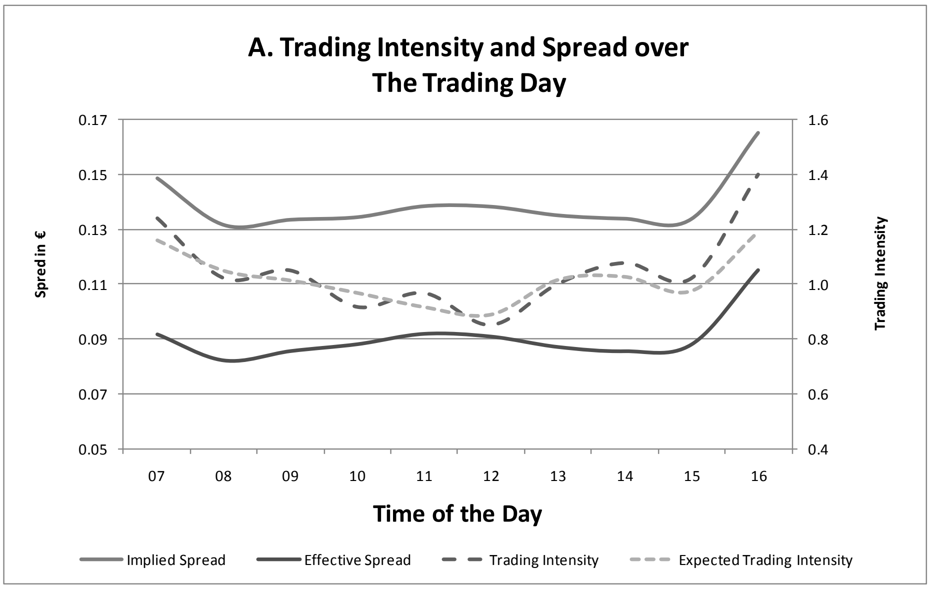
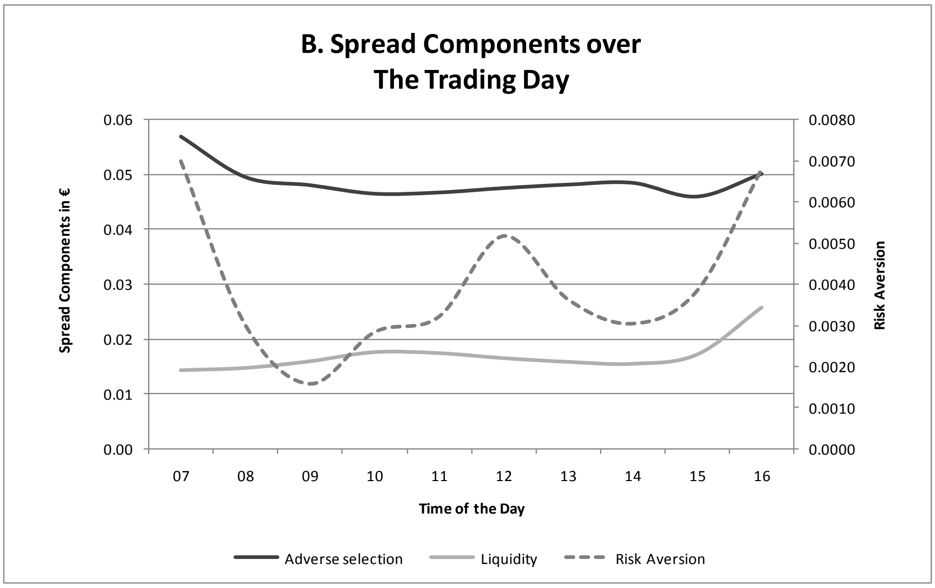
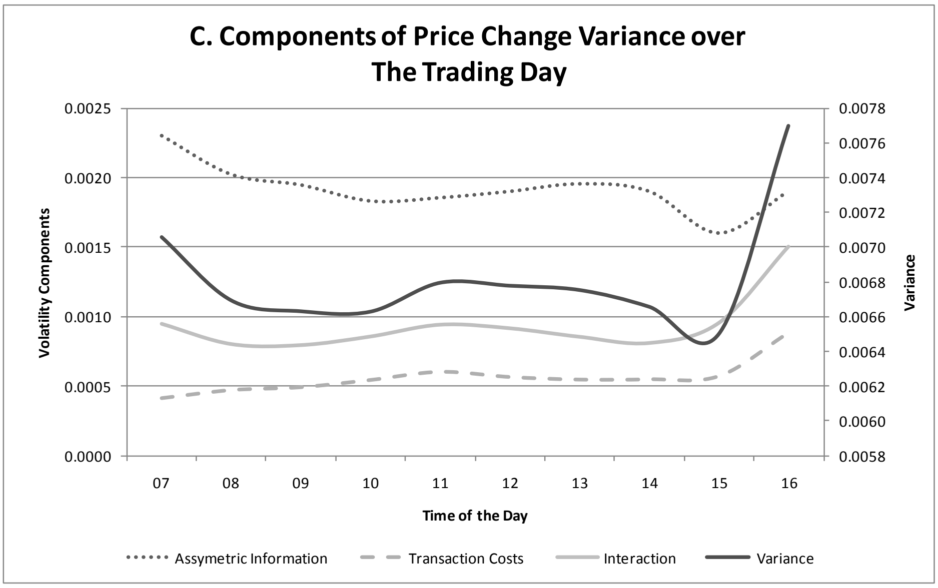


Figure 6.9: ECX II Intraday Variations of Spread, Variance and Their Components

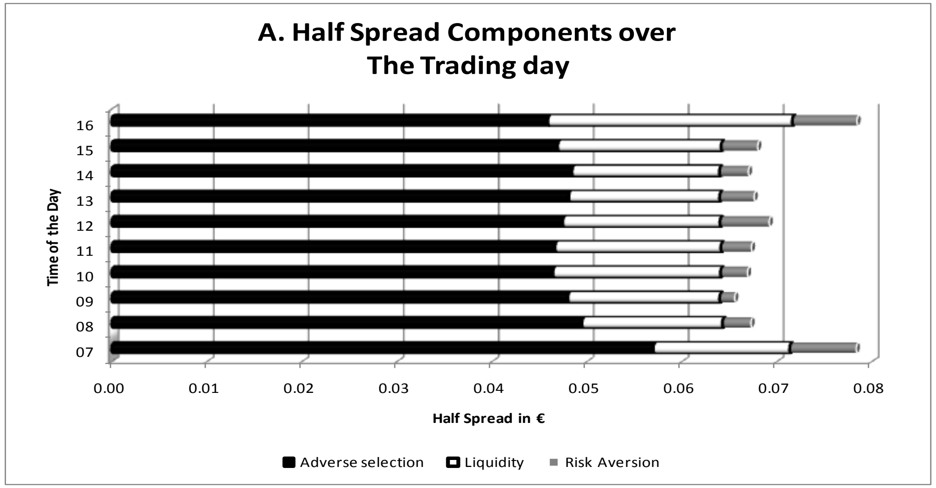


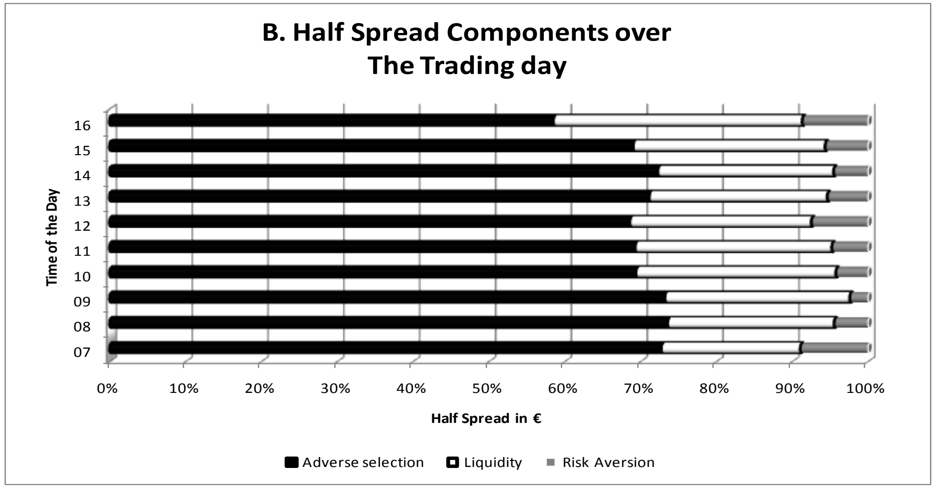


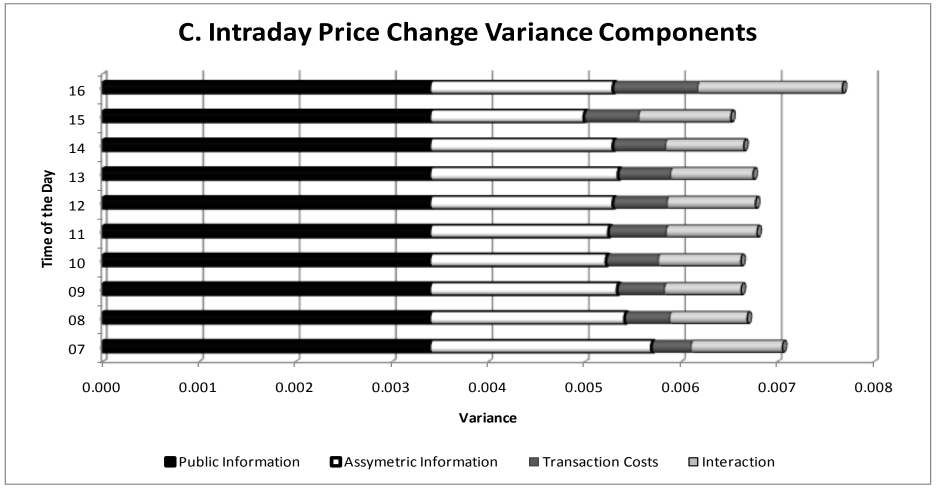


Appendix 6.C

Figure 6.10: ECX II Intraday Spread and Variance Components







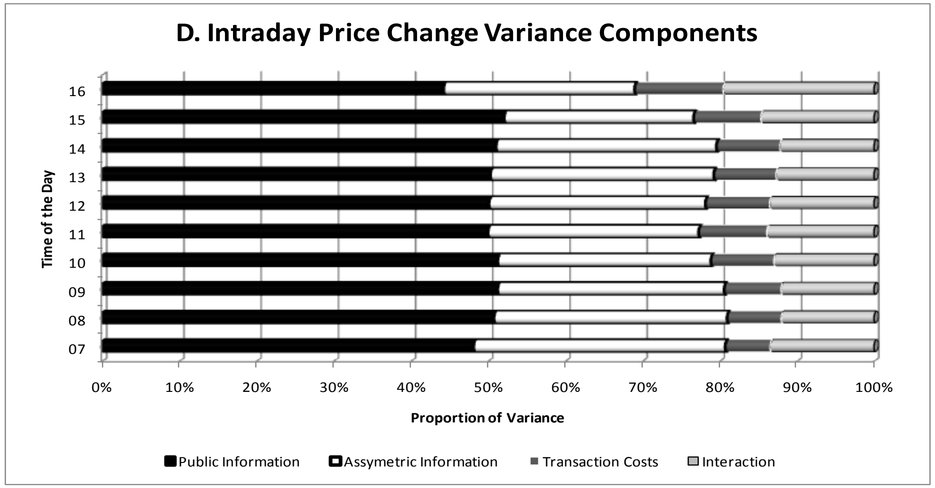
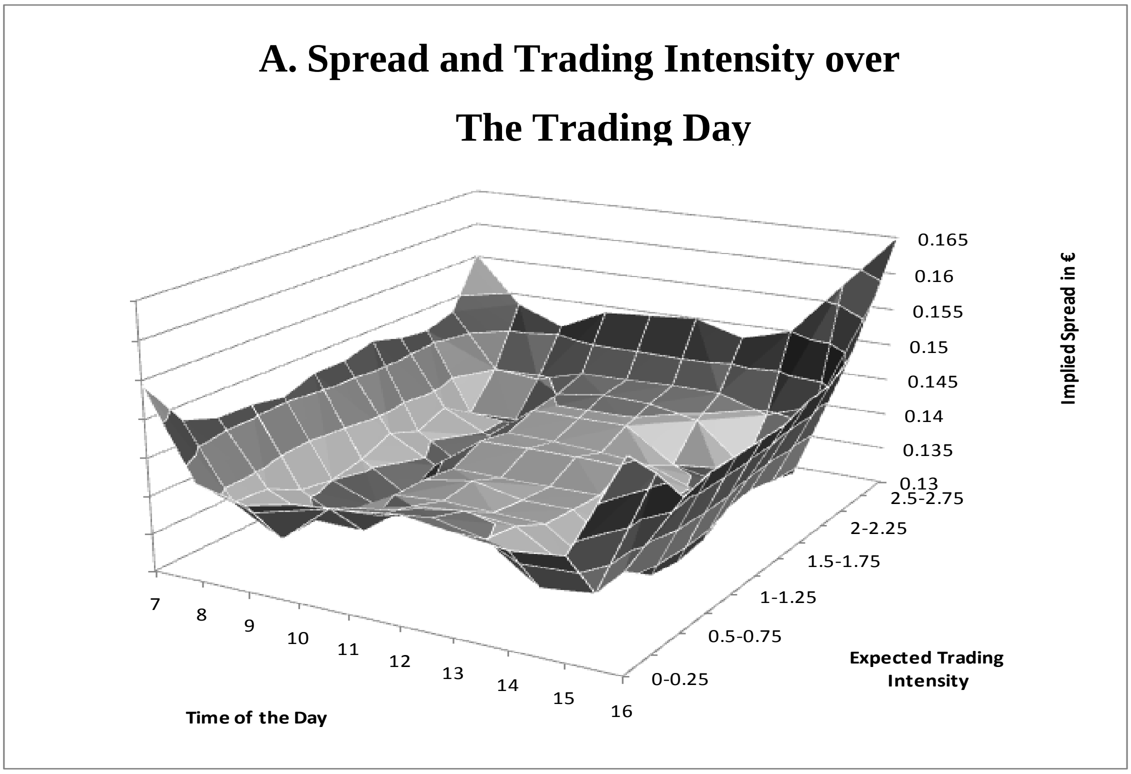


Figure 6.11: ECX II Spread and Variance over the Trading Day and across Trading Intensity



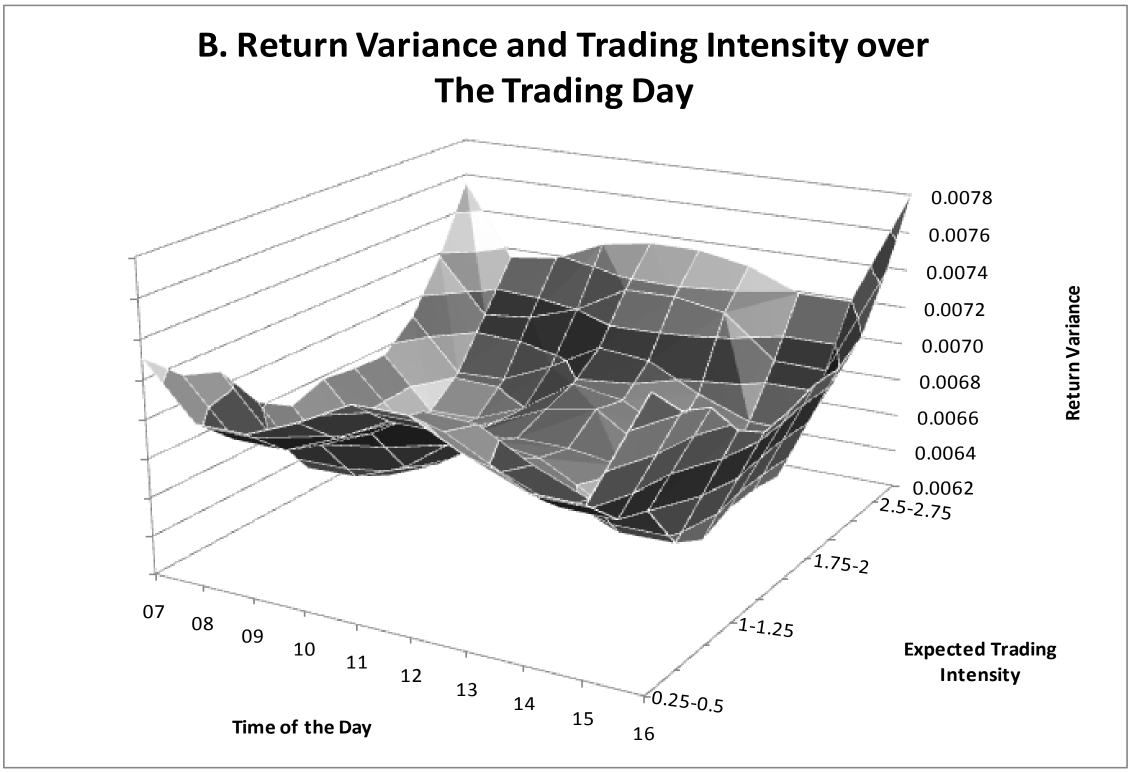
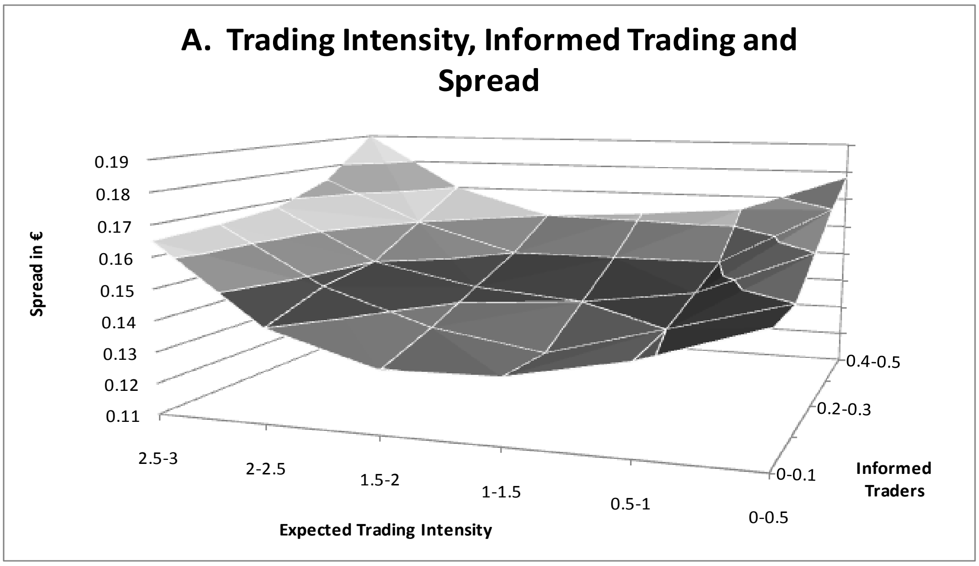
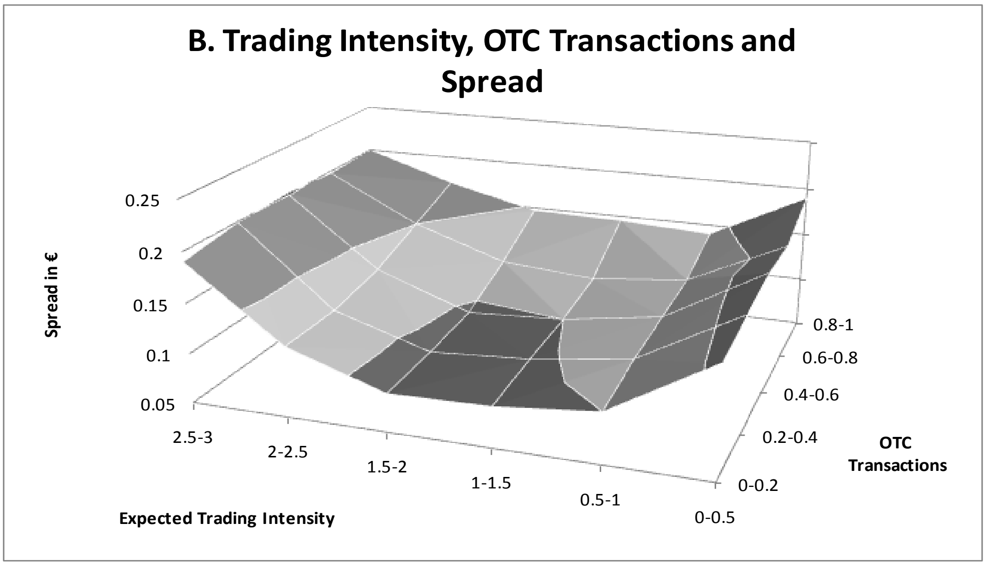


Figure 6.12: ECX II Spread across Trading Intensity, Risk and OTC Transactions





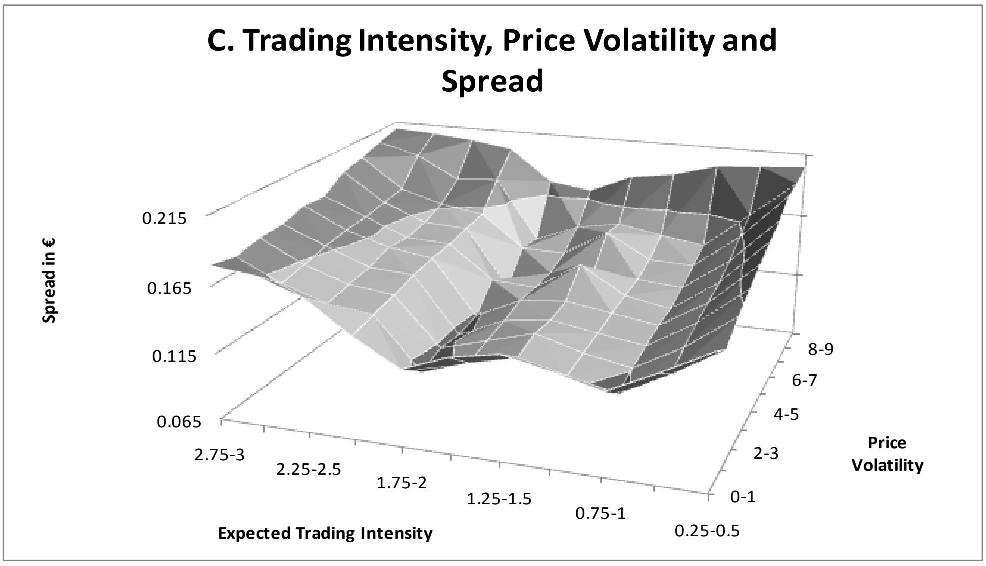
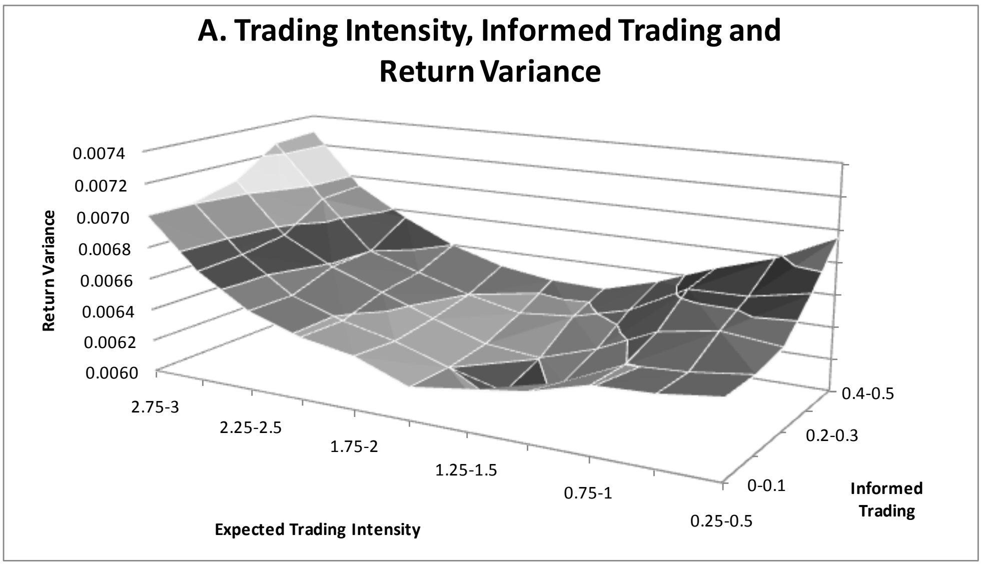
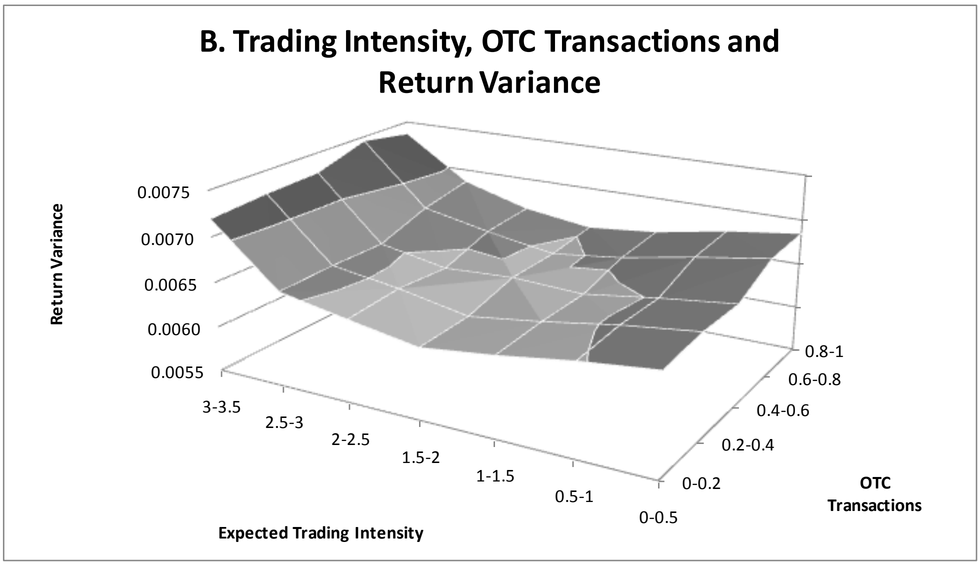


Figure 6.13: ECX II Variance across Trading Intensity, Risk and OTC Transactions





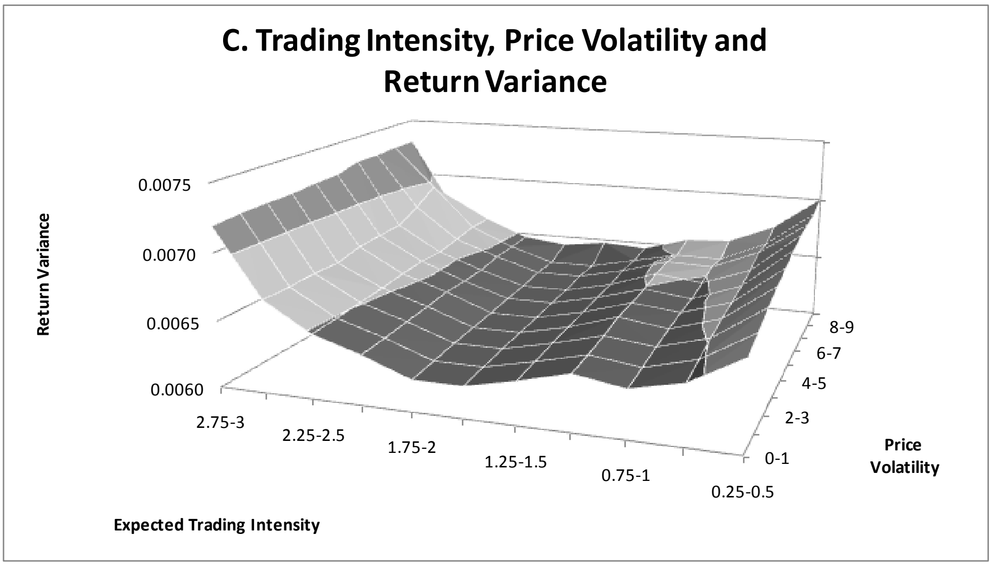
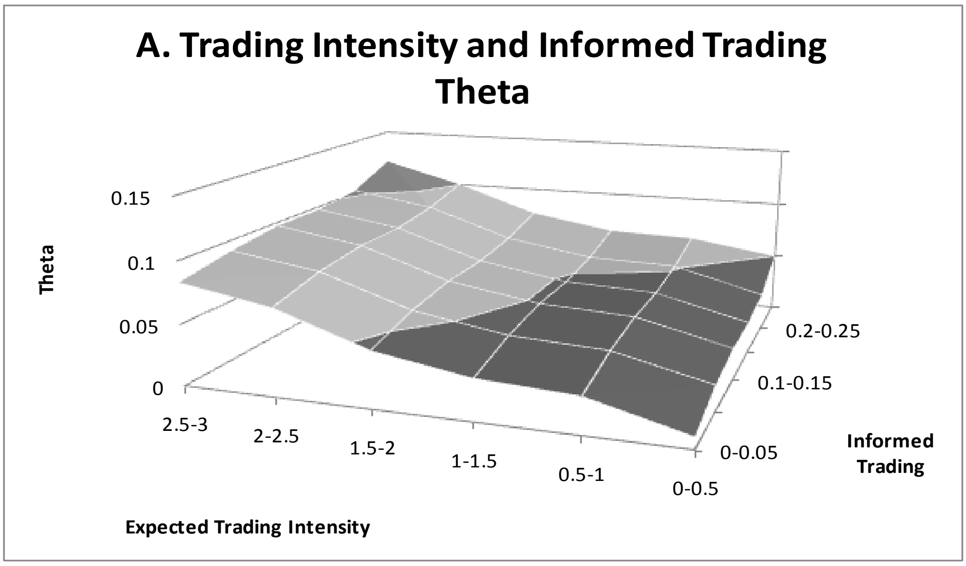
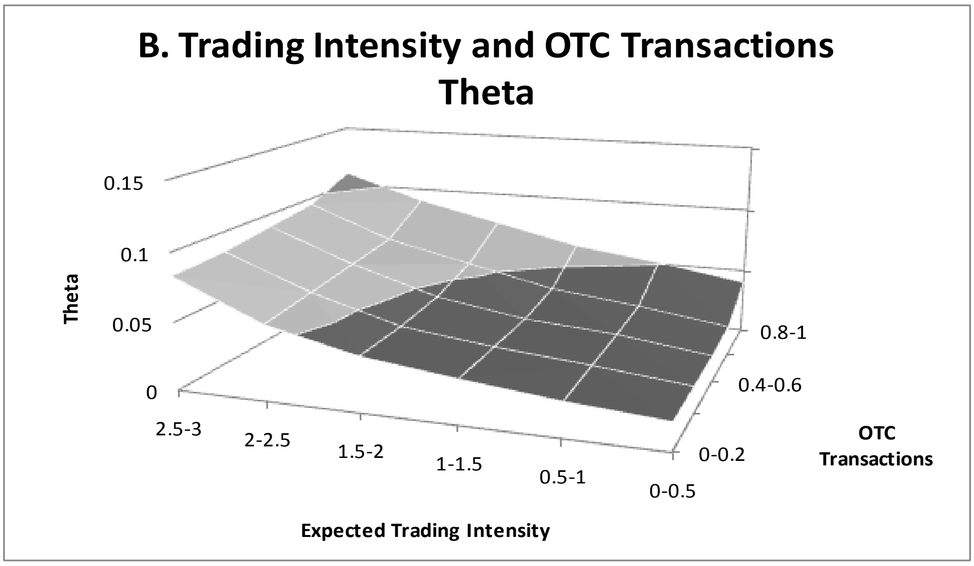


Figure 6.14: ECX II Theta across Trading Intensity, Risk and OTC Transactions





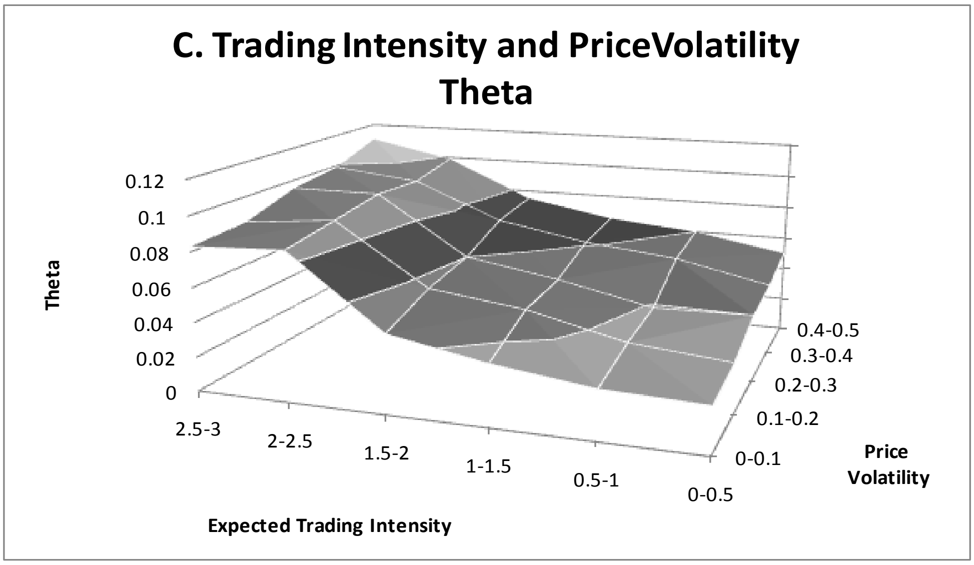
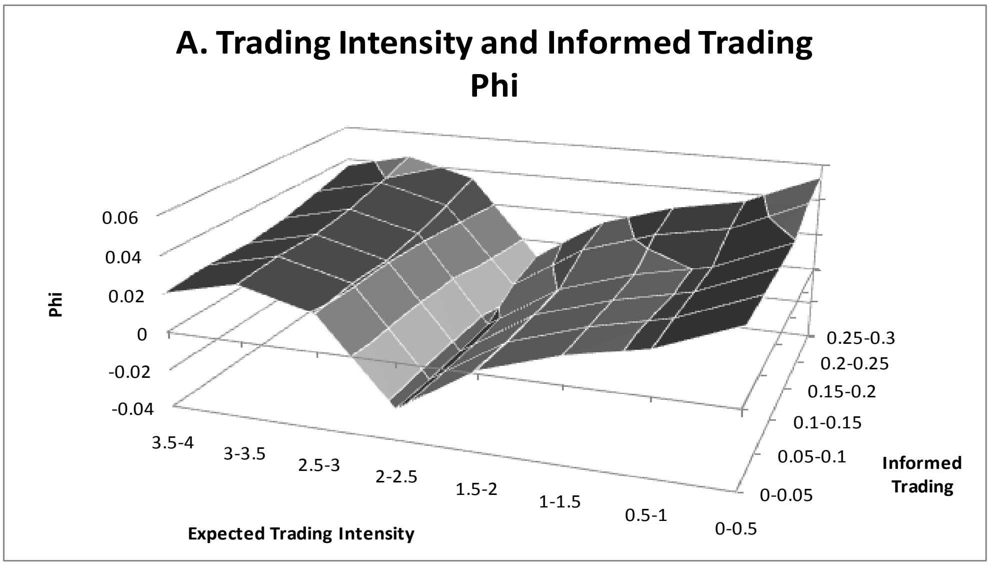
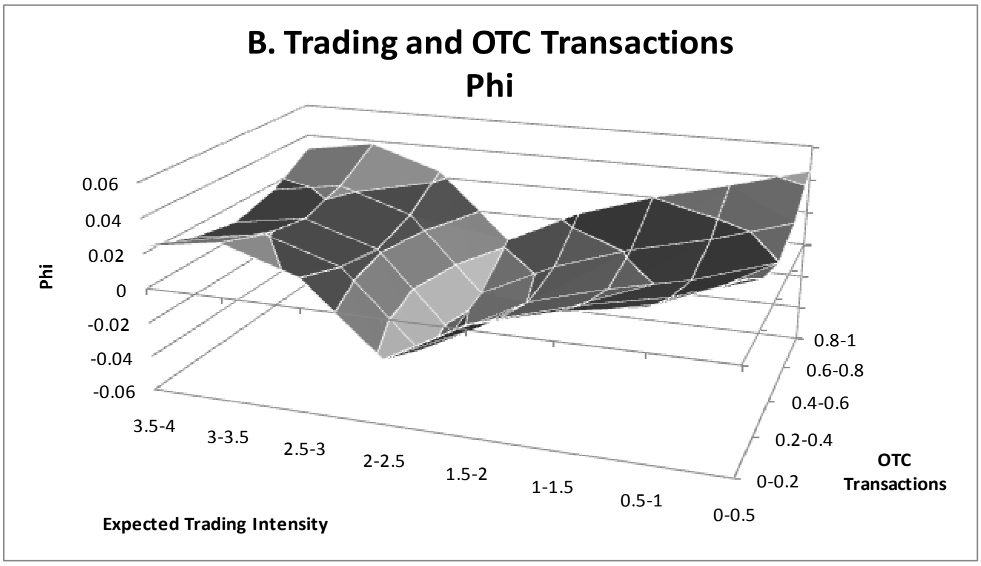


Figure 6.15: ECX II Phi across Trading Intensity, Risk and OTC Transactions





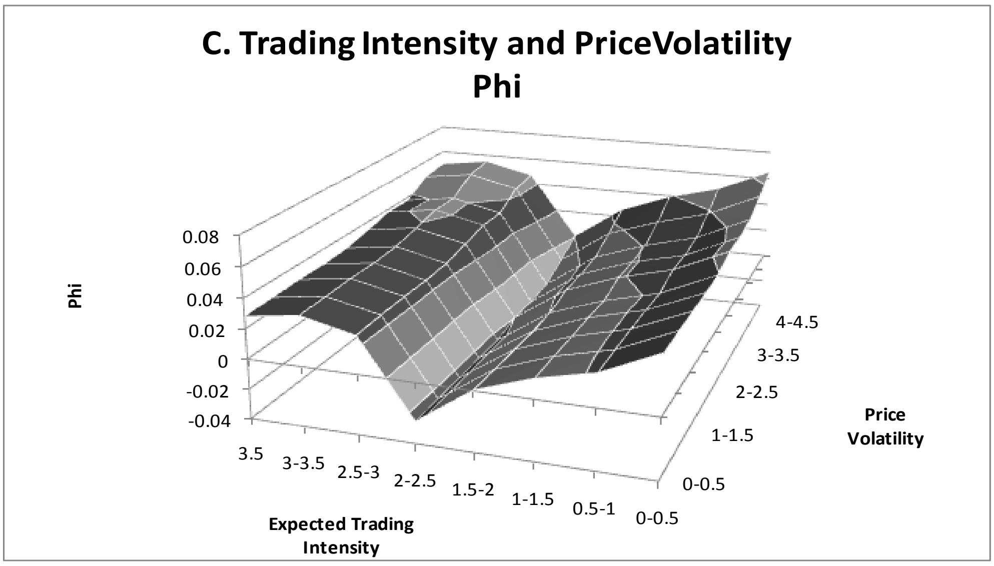
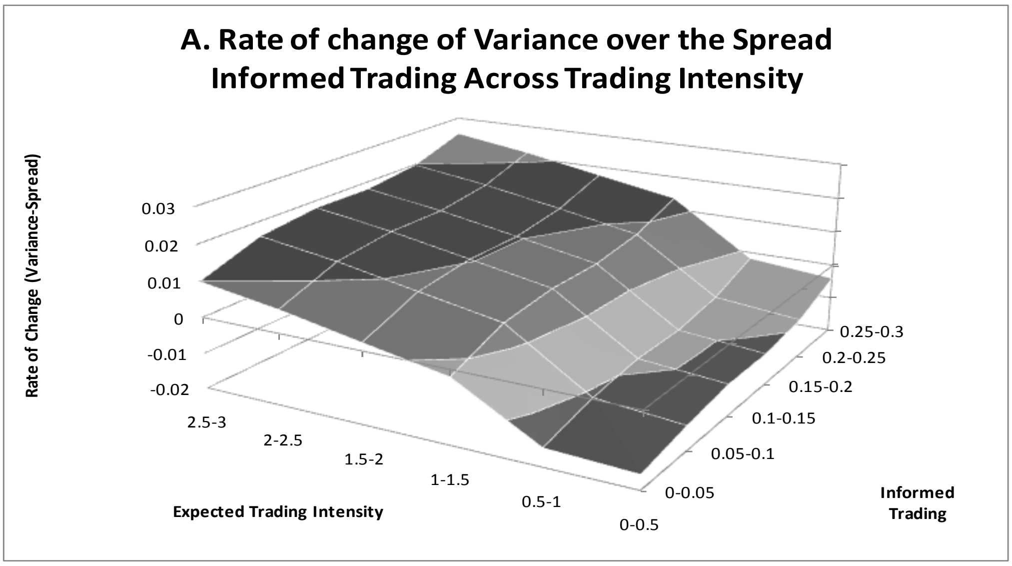
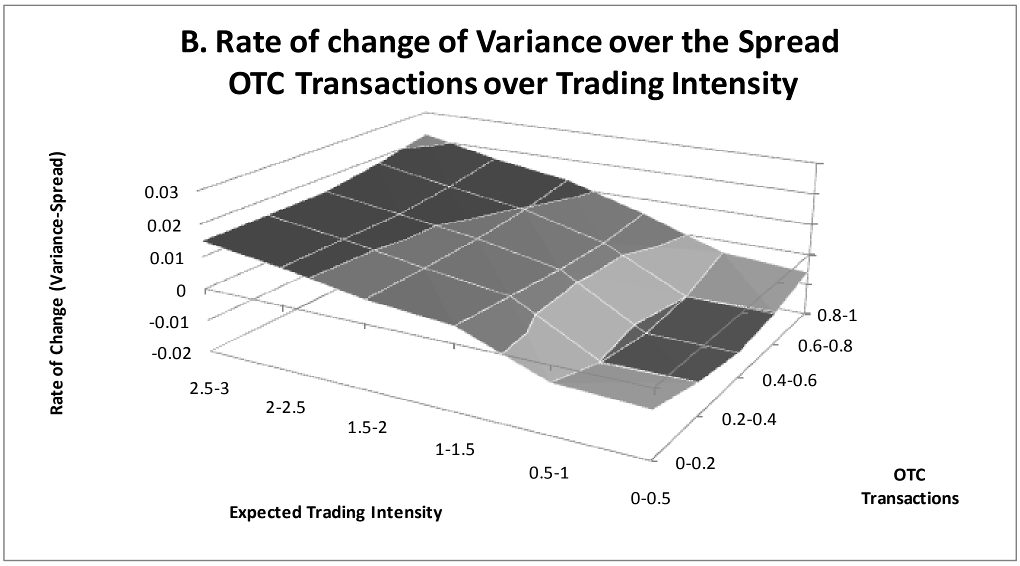


Figure 6.16: ECX II Marginal Variance over Spread change across Trading Intensity, Risk and OTC Transactions





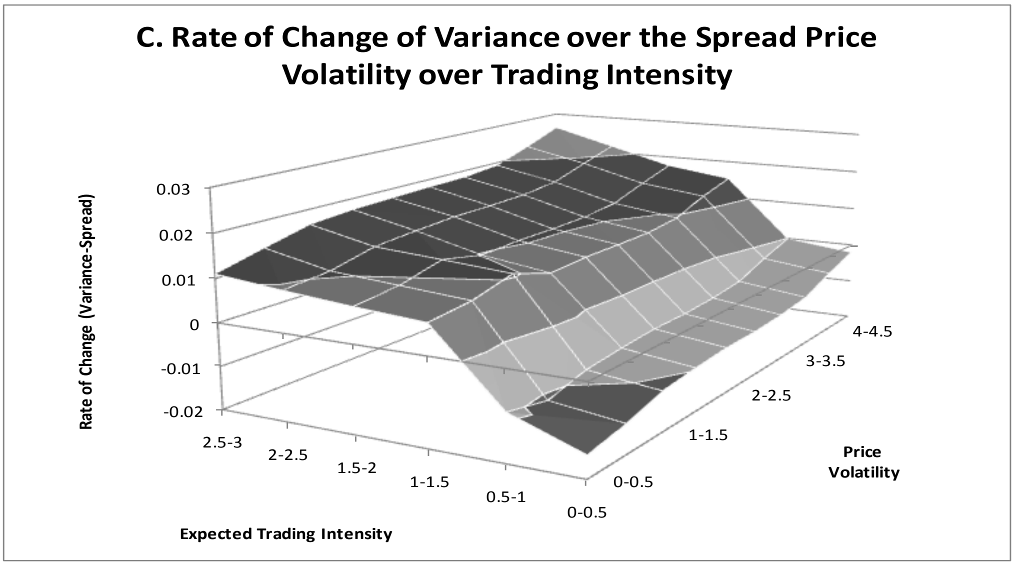
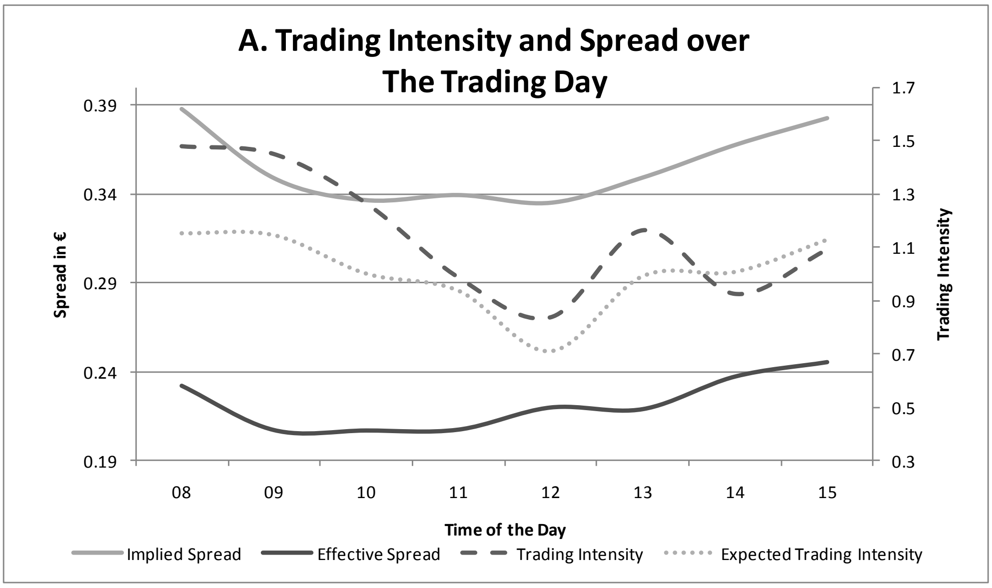
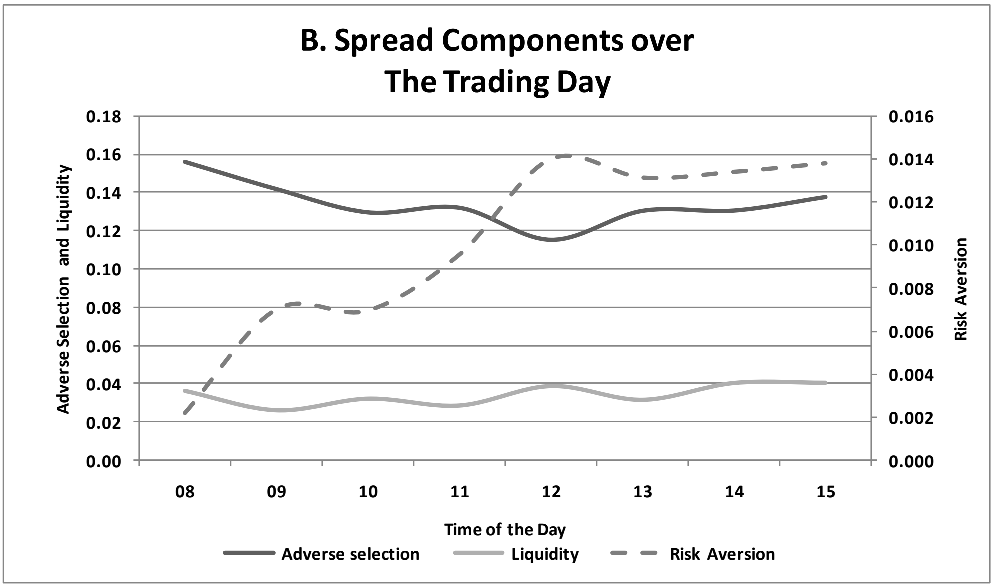
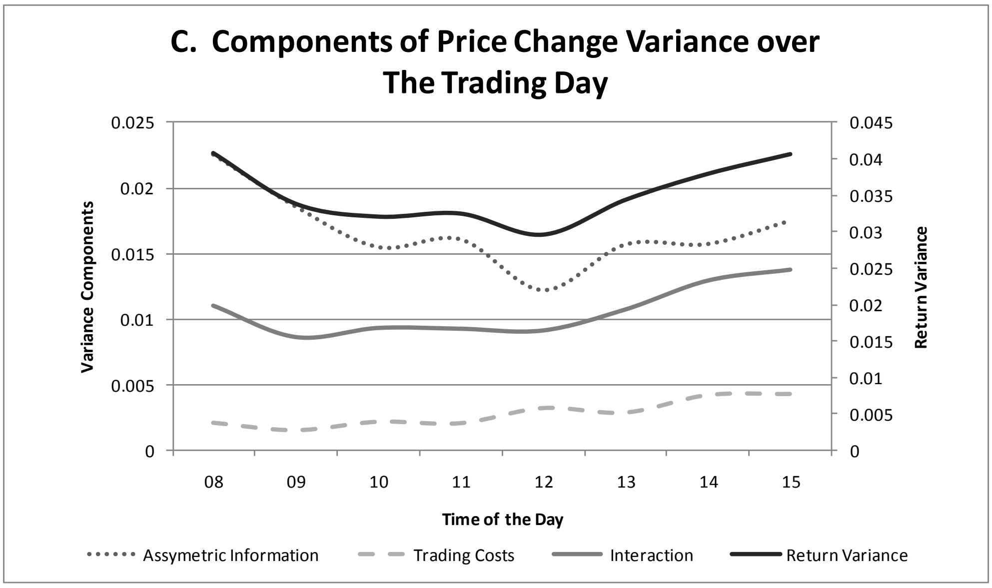


Figure 6.17: NP I Intraday Variations of Spread, Variance and Their Components







Appendix 6.C

Figure 6.18: NP I Intraday Spread and Variance Components

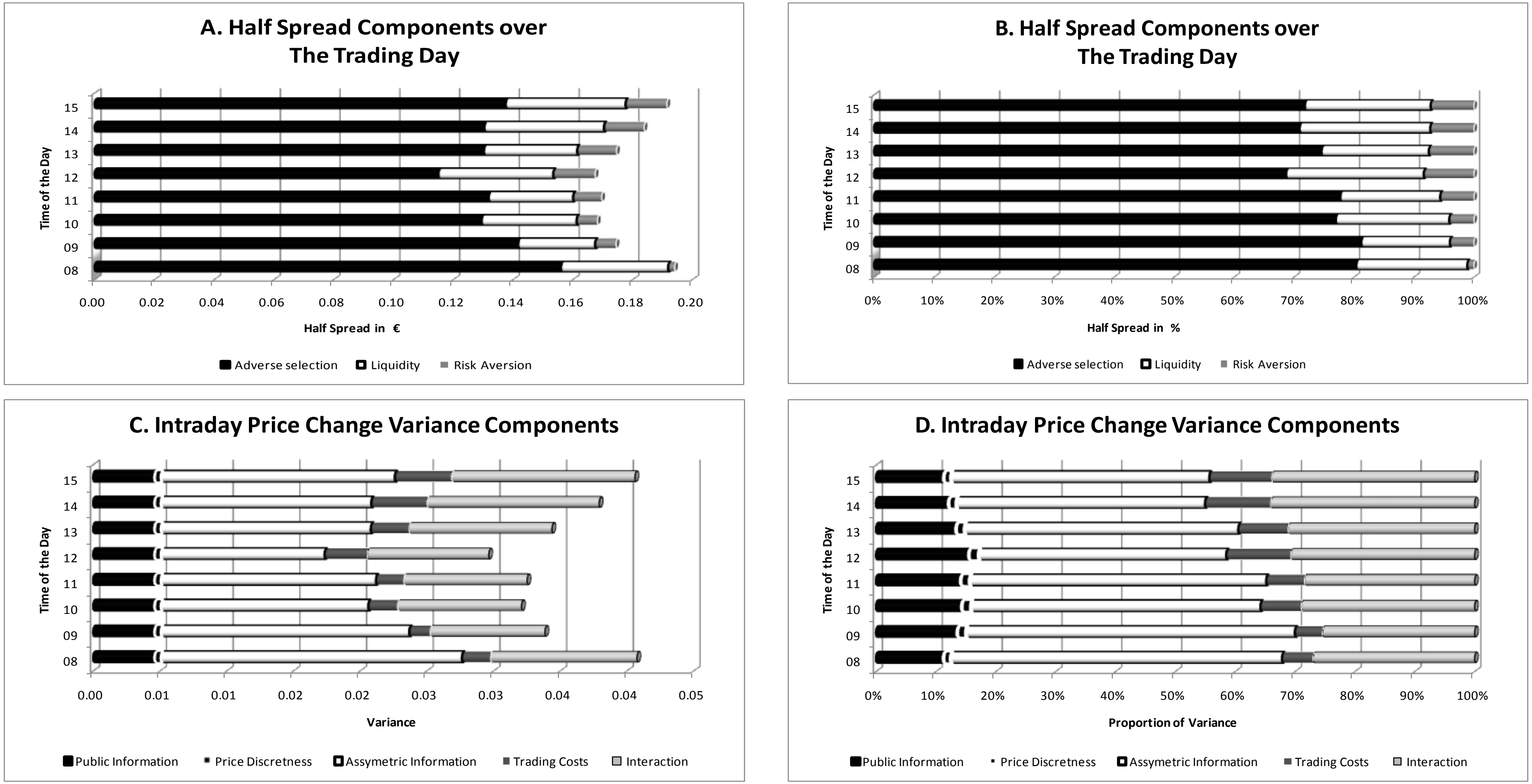
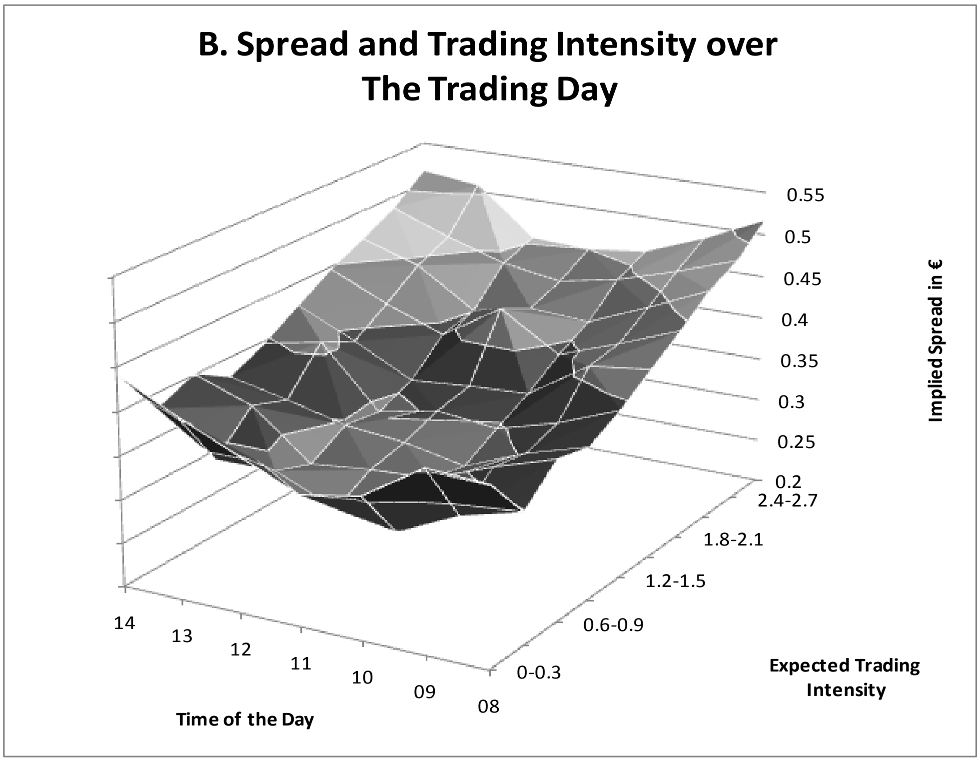


Figure 6.19: NP I Spread and Variance over the Trading Day and across Trading Intensity



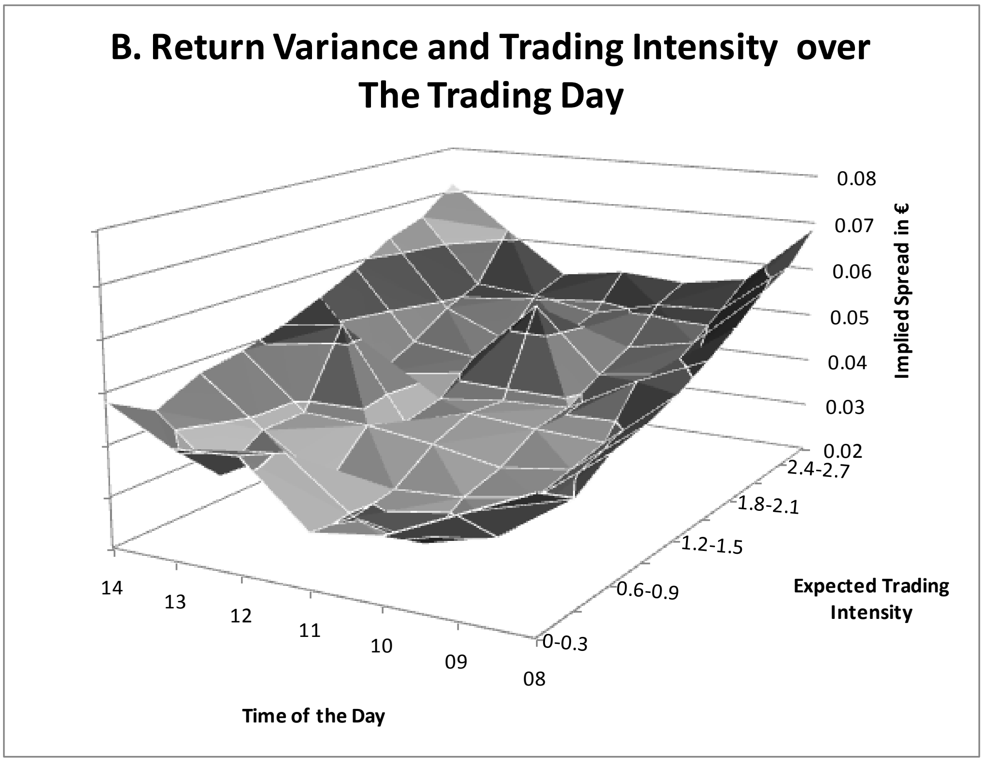
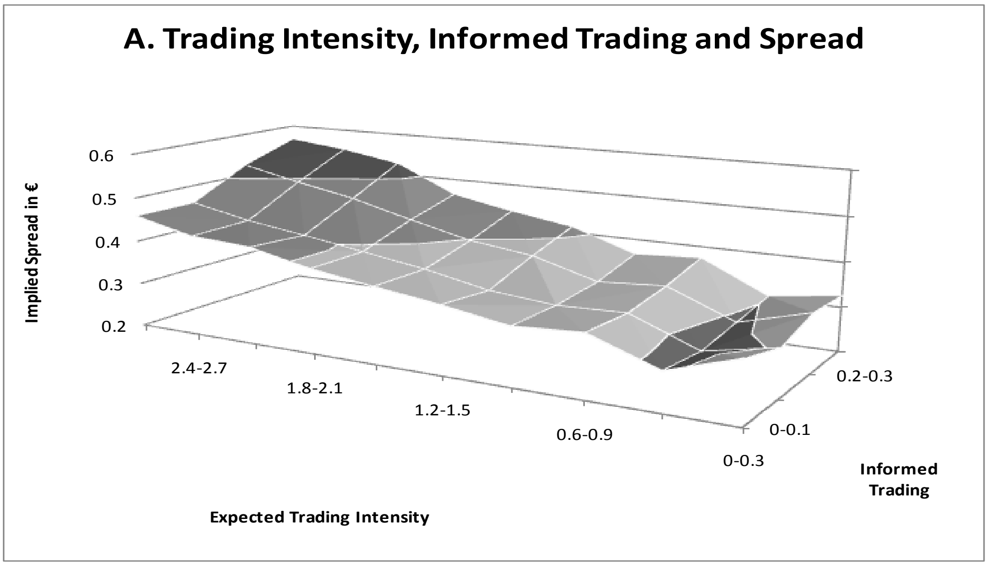
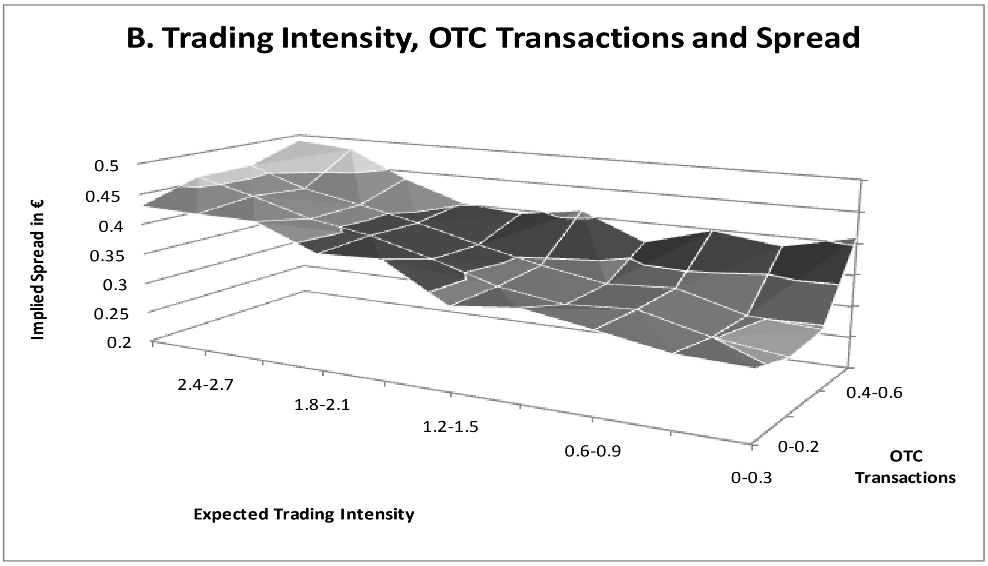


Figure 6.20: NP I Spread across Trading Intensity, Risk and OTC Transactions





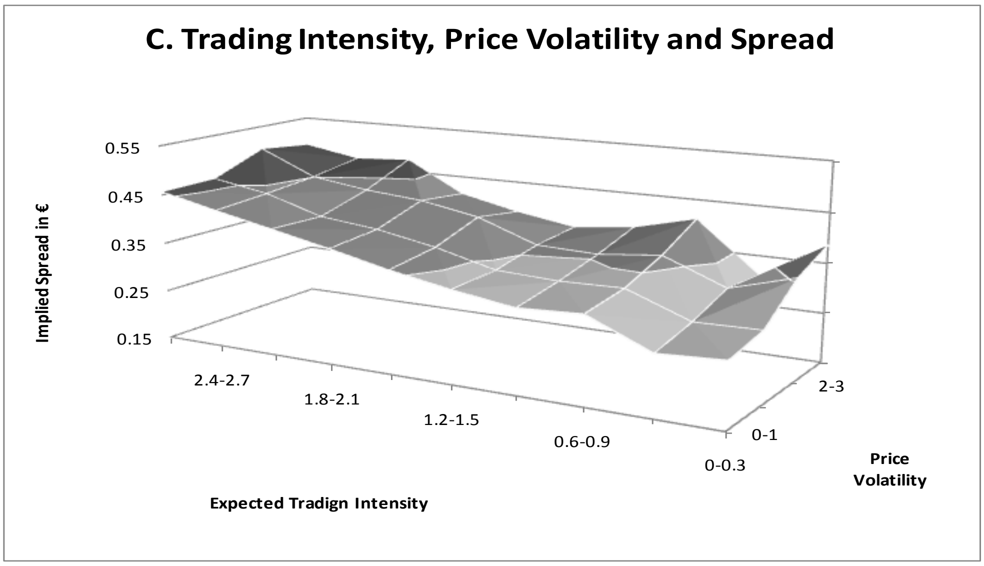
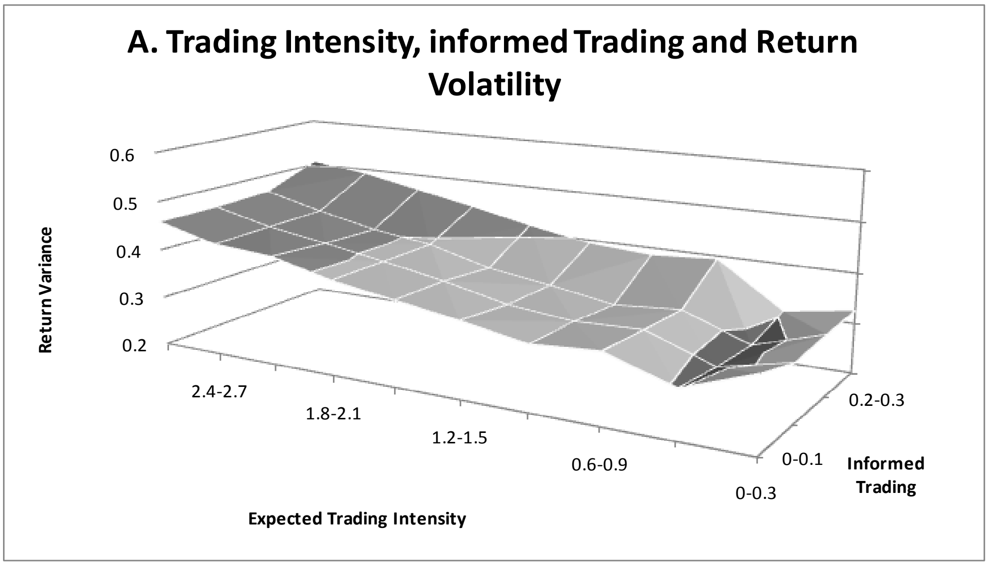
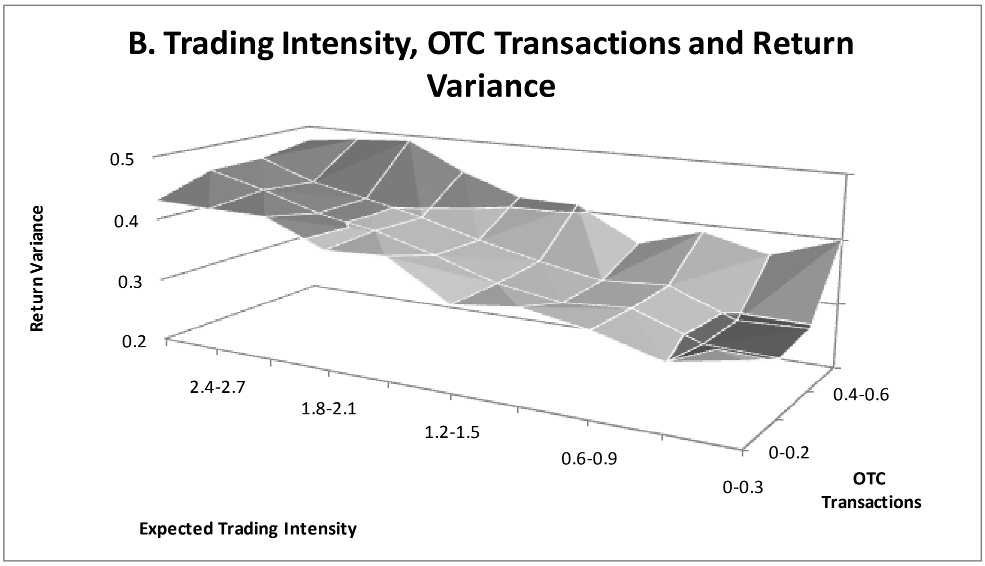


Figure 6.21: NP I Variance across Trading Intensity, Risk and OTC Transactions





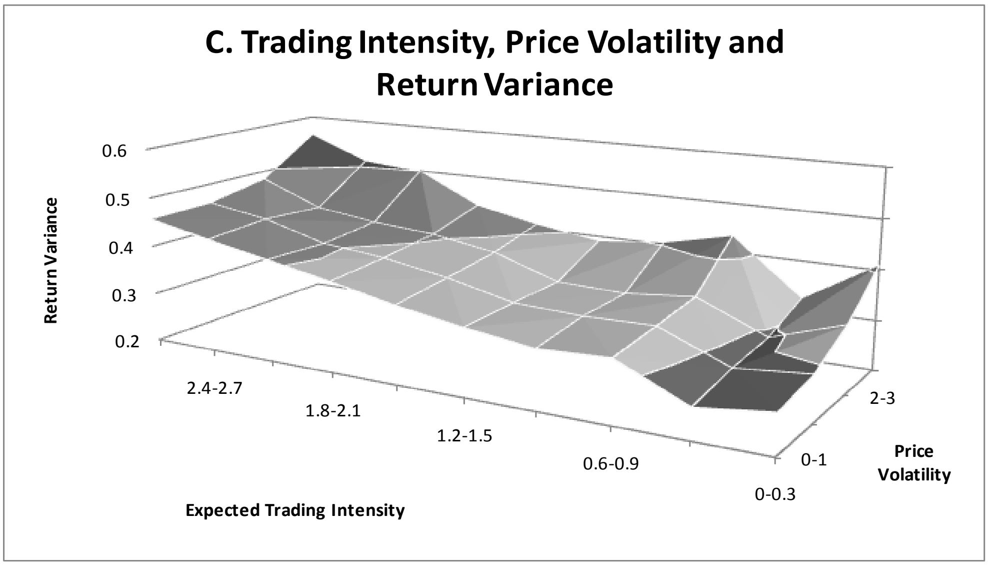
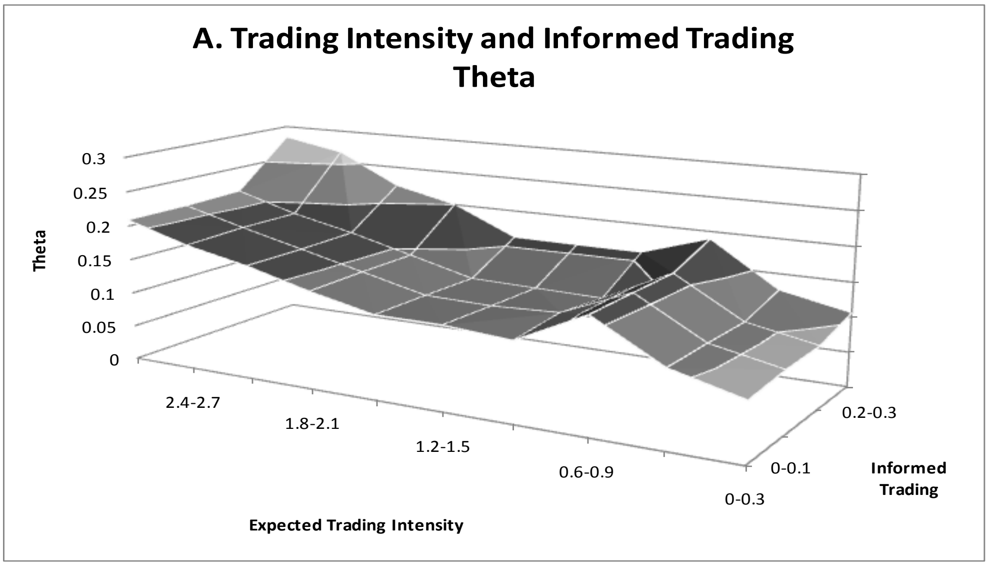
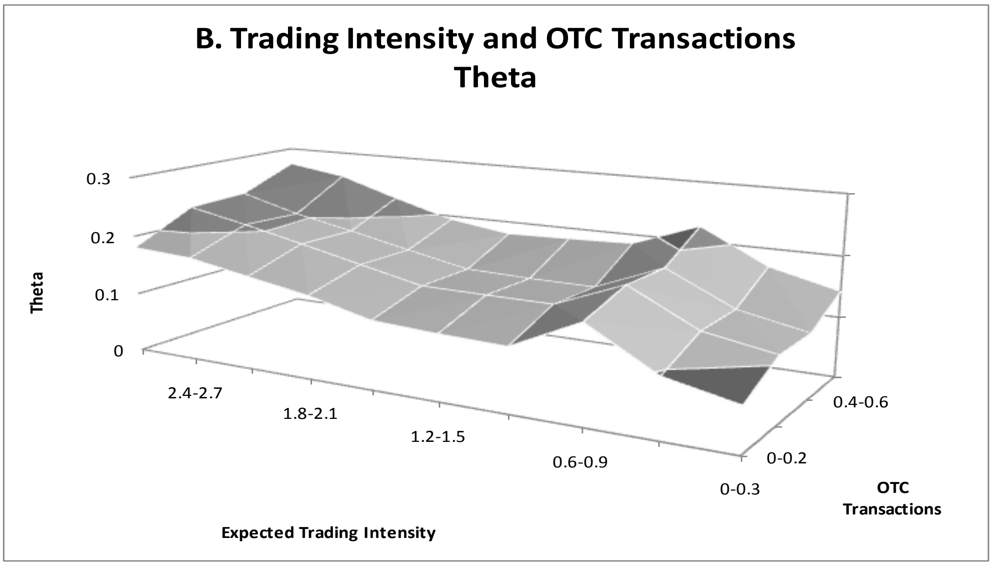


Figure 6.22: NP I Theta across Trading Intensity, Risk and OTC Transactions





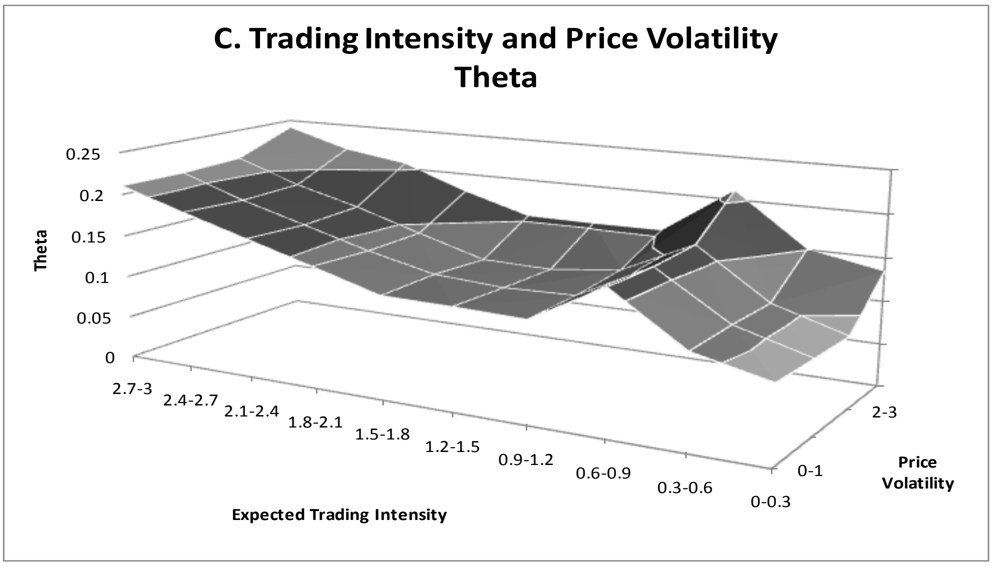
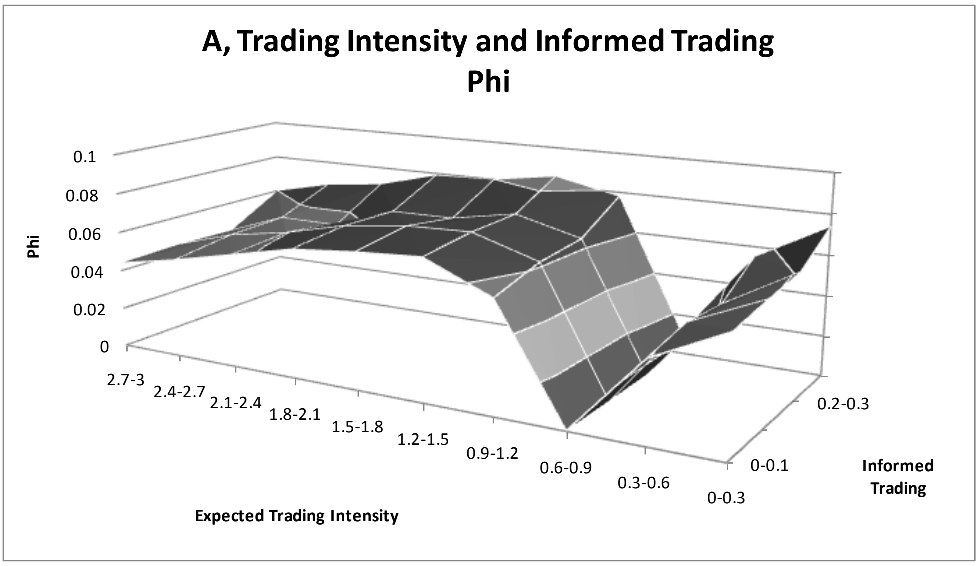
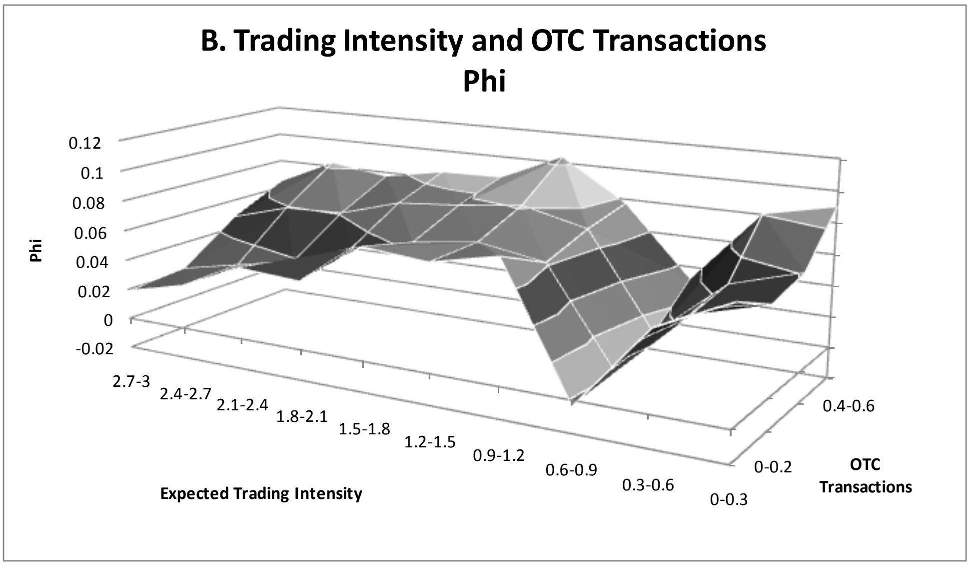


Figure 6.23: NP I Phi across Trading Intensity, Risk and OTC Transactions





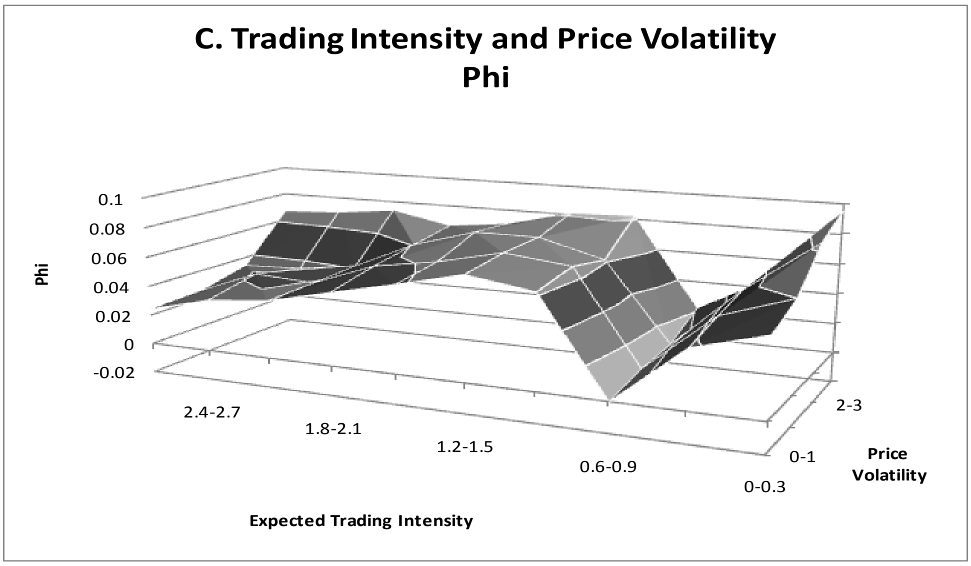
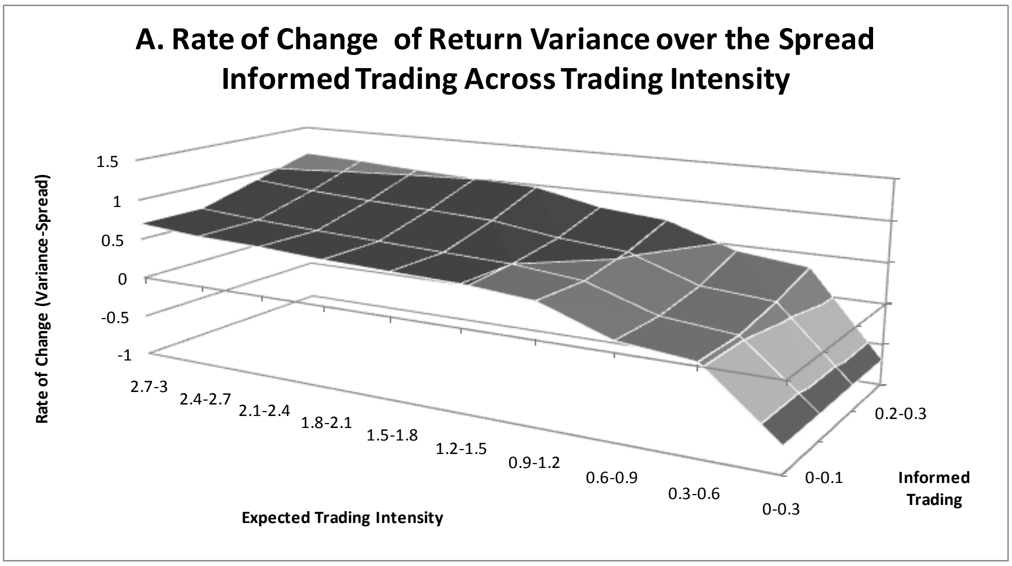
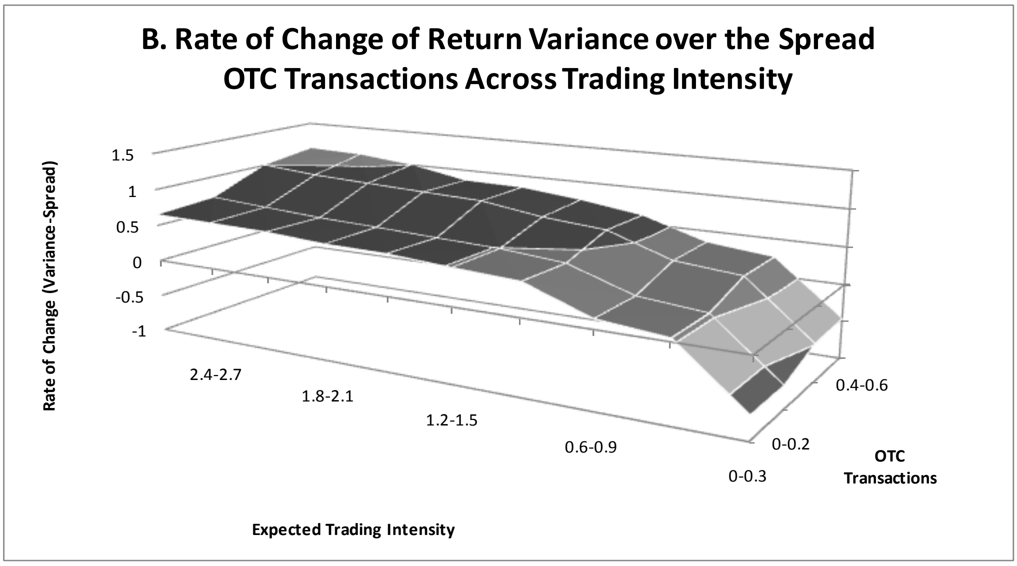


Figure 6.24: NP I Marginal Variance over Spread change across Trading Intensity, Risk and OTC Transactions





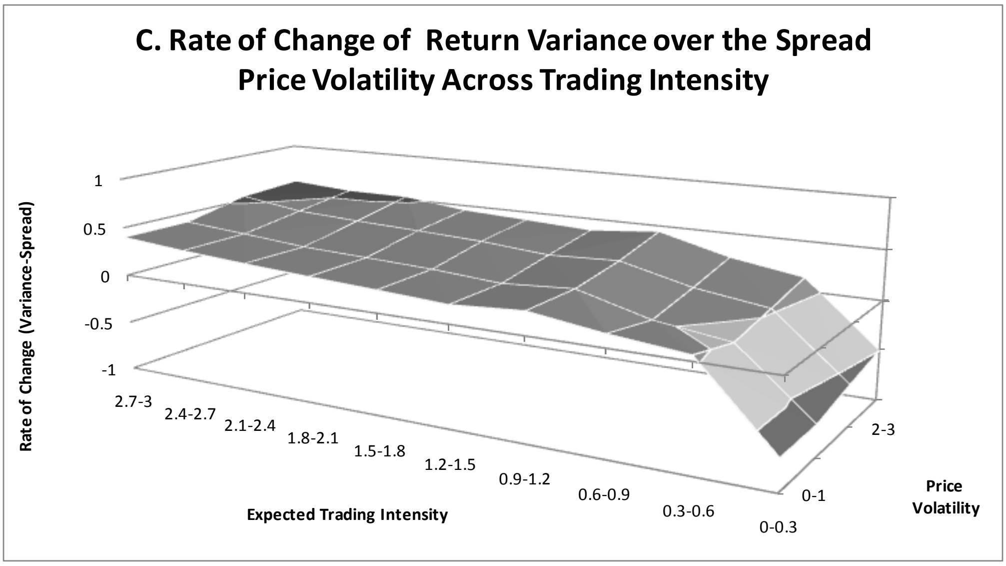


Figure 6.25: NP II Intraday Variations of Spread, Variance and Their Components

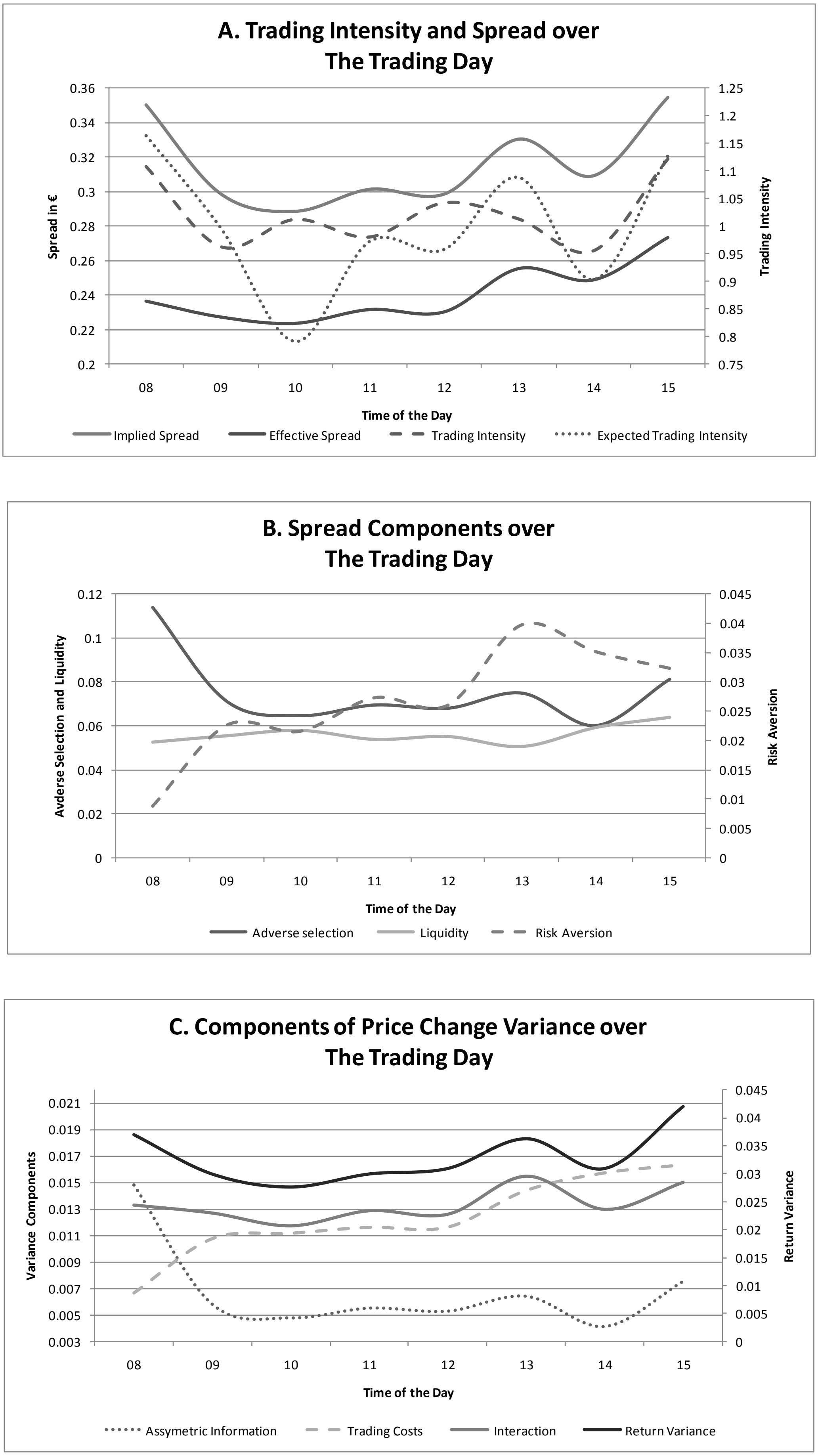


Figure 6.26: NP II Intraday Spread and Variance Components

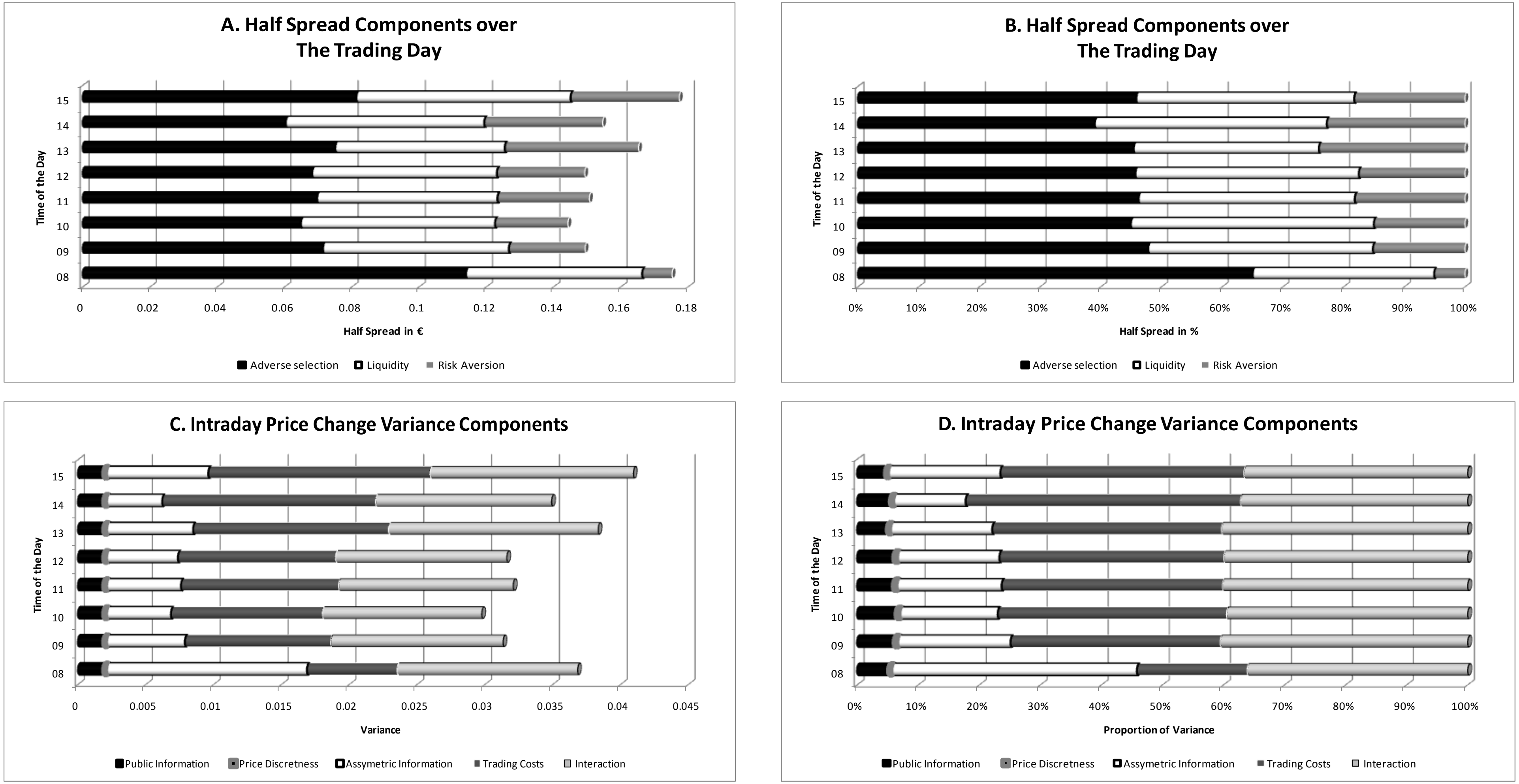
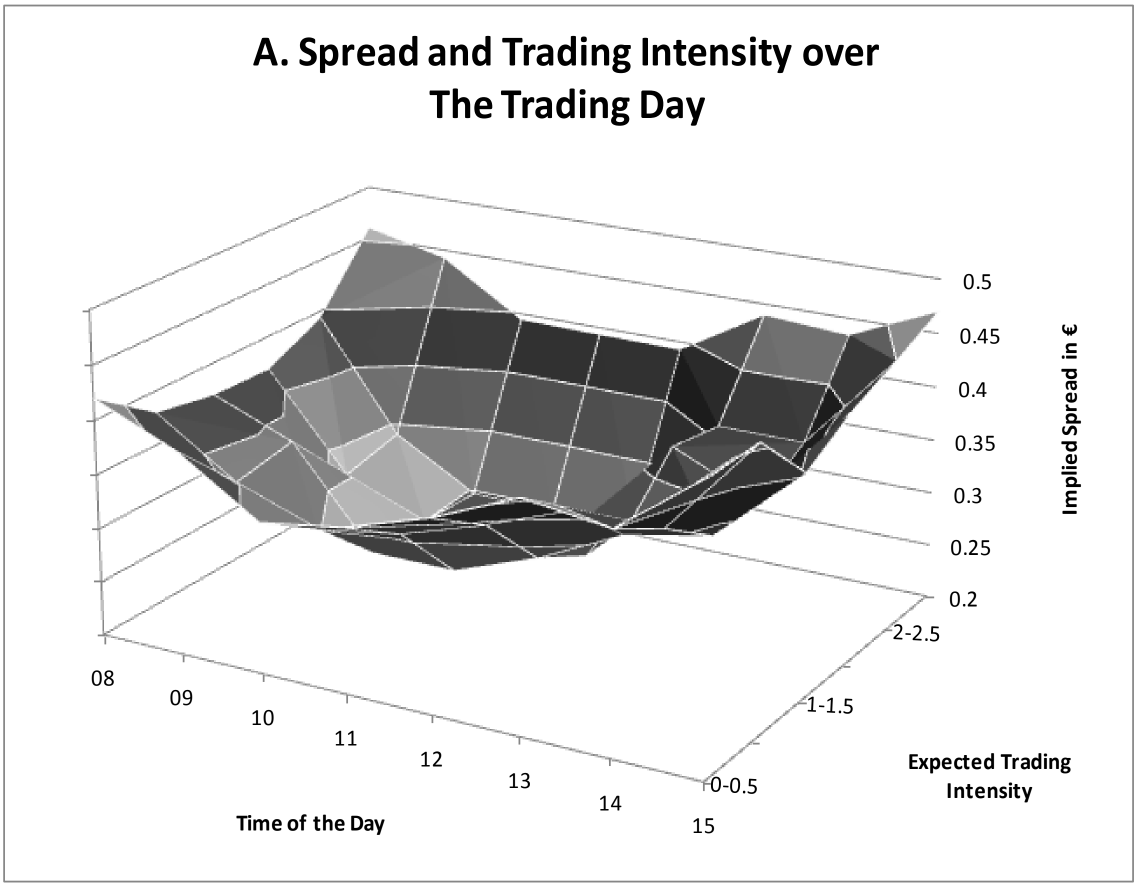


Figure 6.27: NP II Spread and Variance over the Trading Day and across Trading Intensity



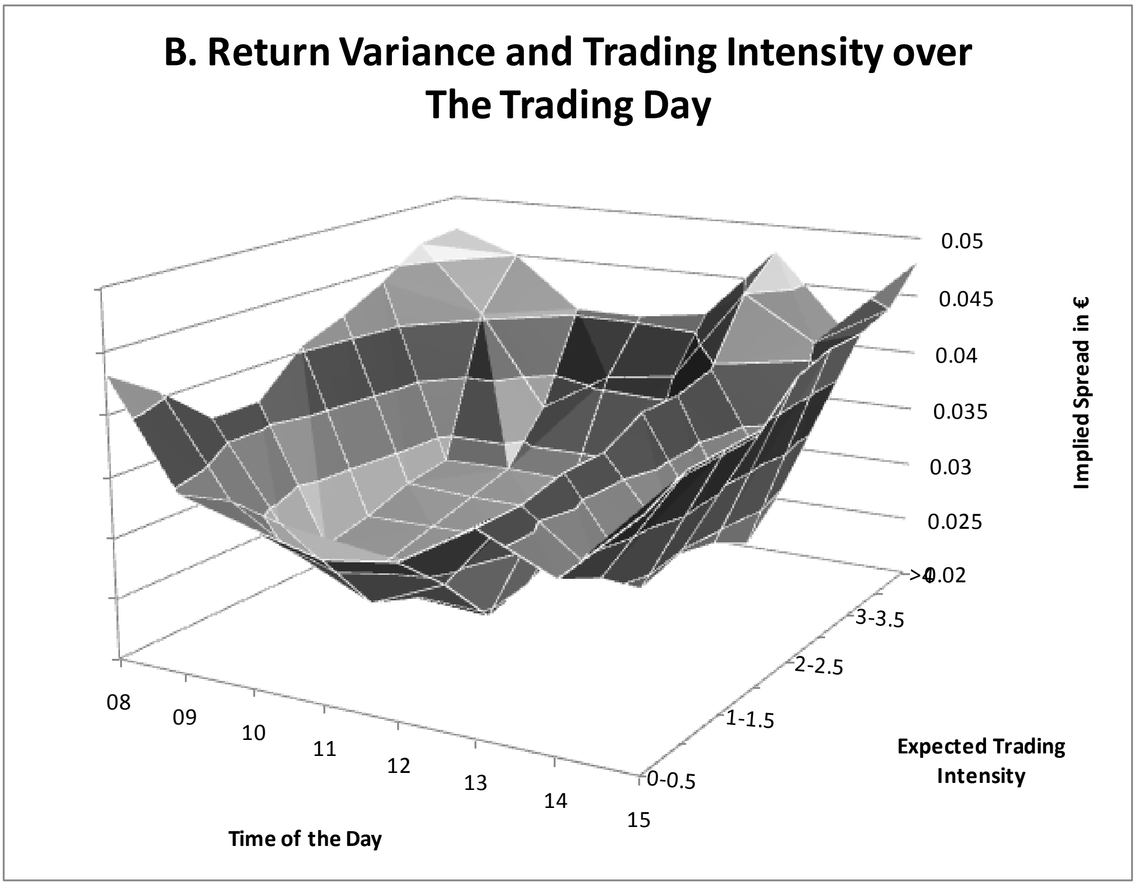
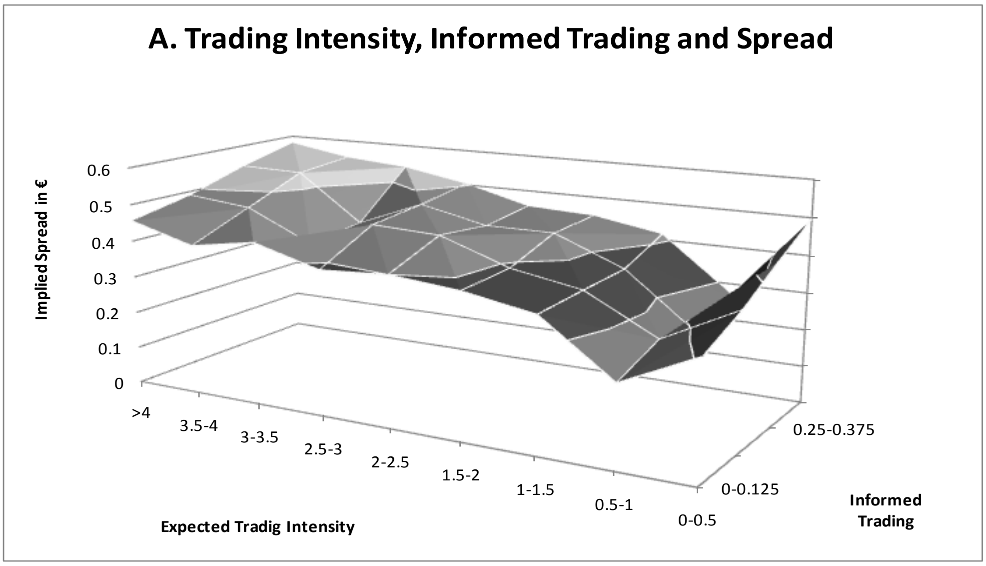
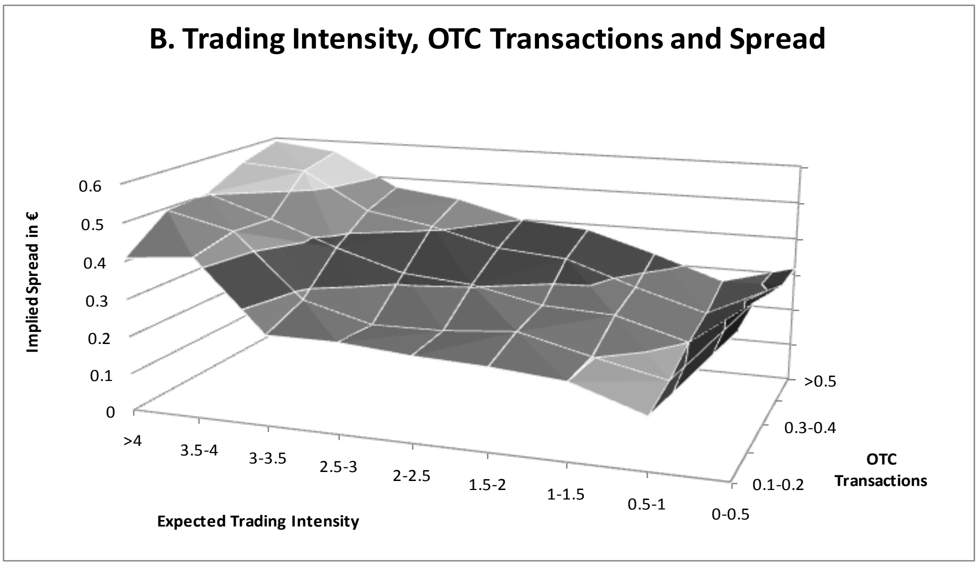


Figure 6.28: NP II Spread across Trading Intensity, Risk and OTC Transactions





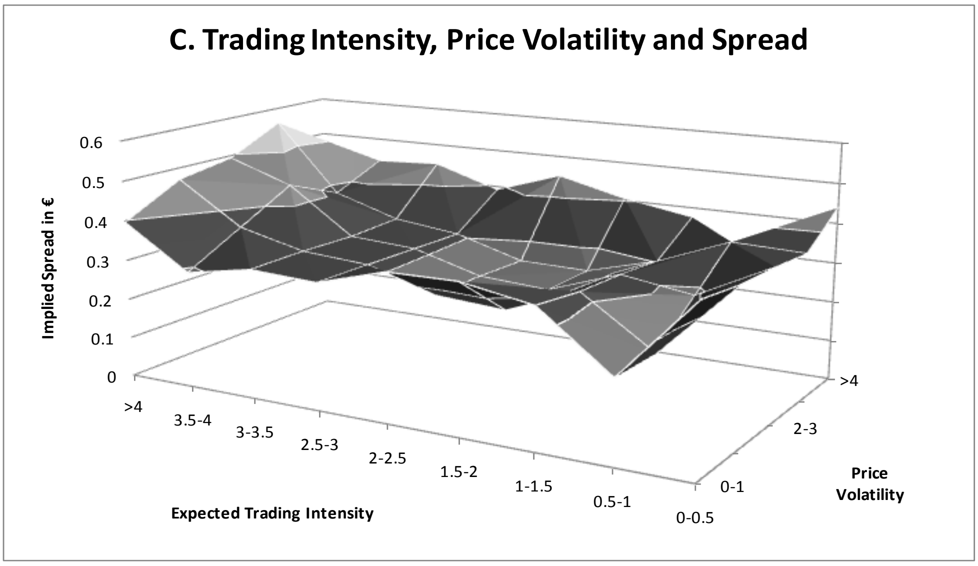
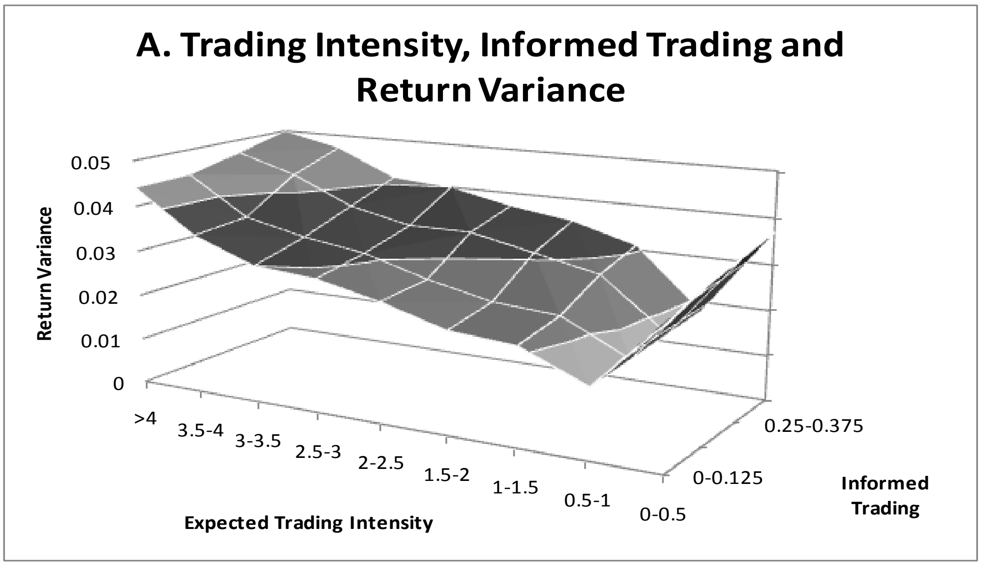
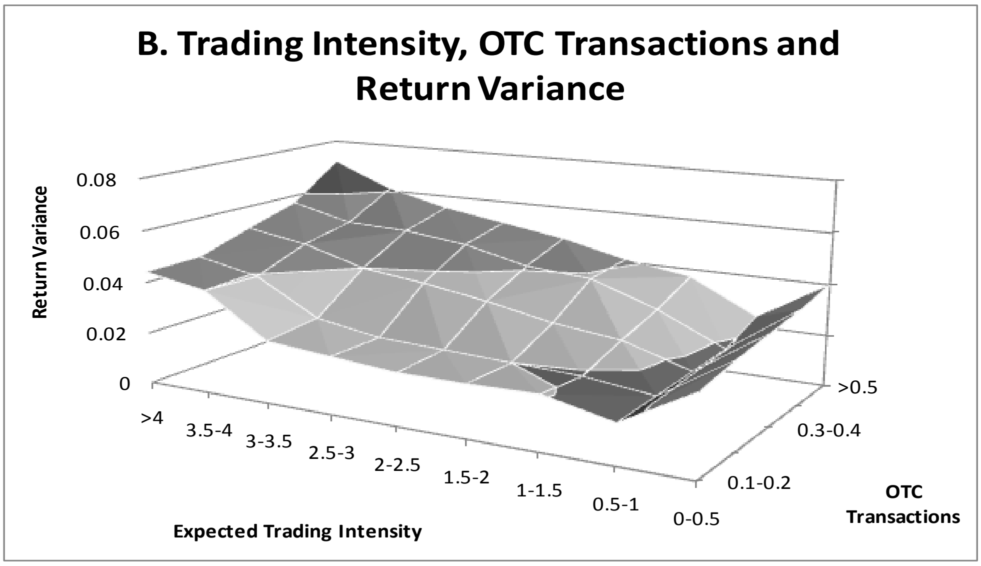


Figure 6.29: NP II Variance across Trading Intensity, Risk and OTC Transactions





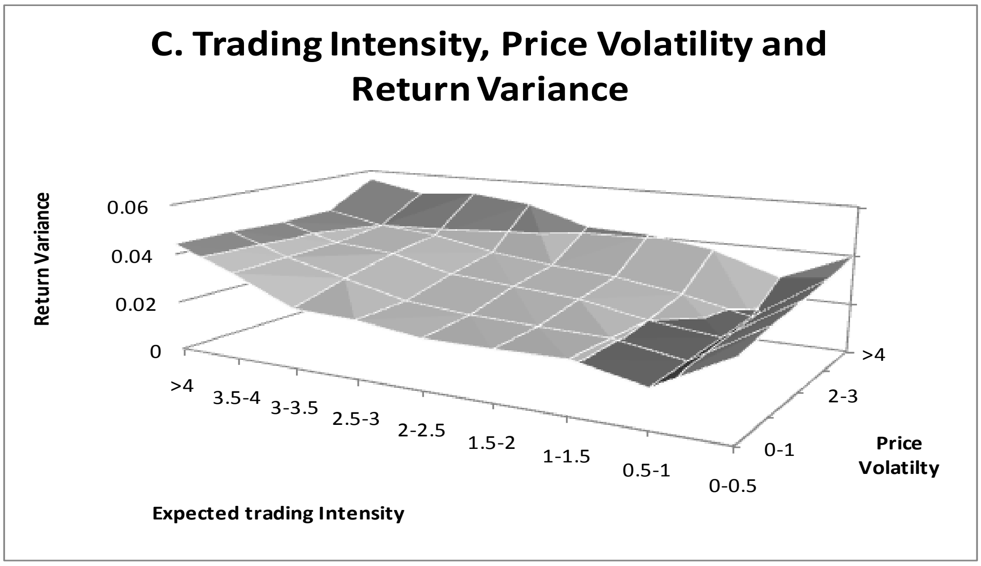
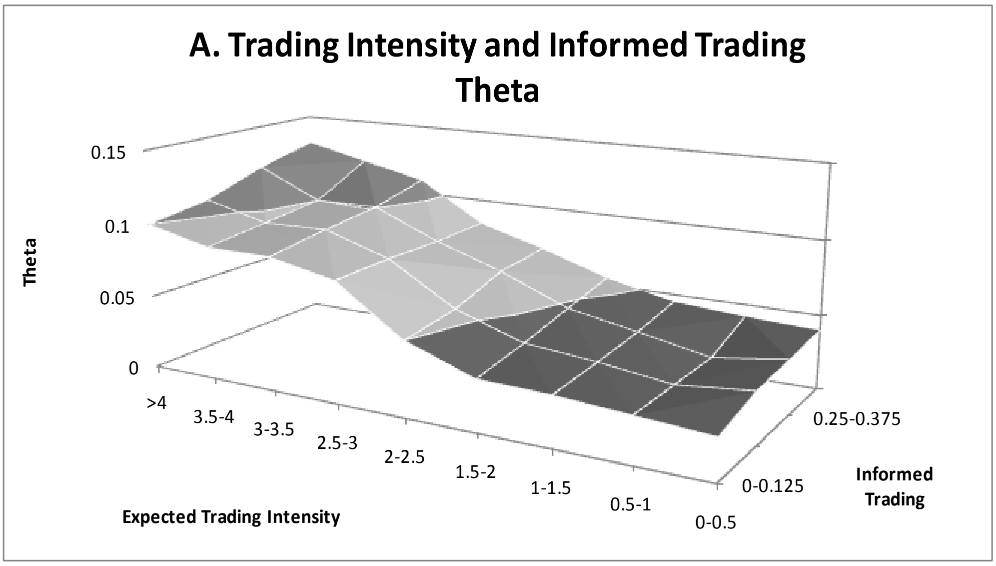
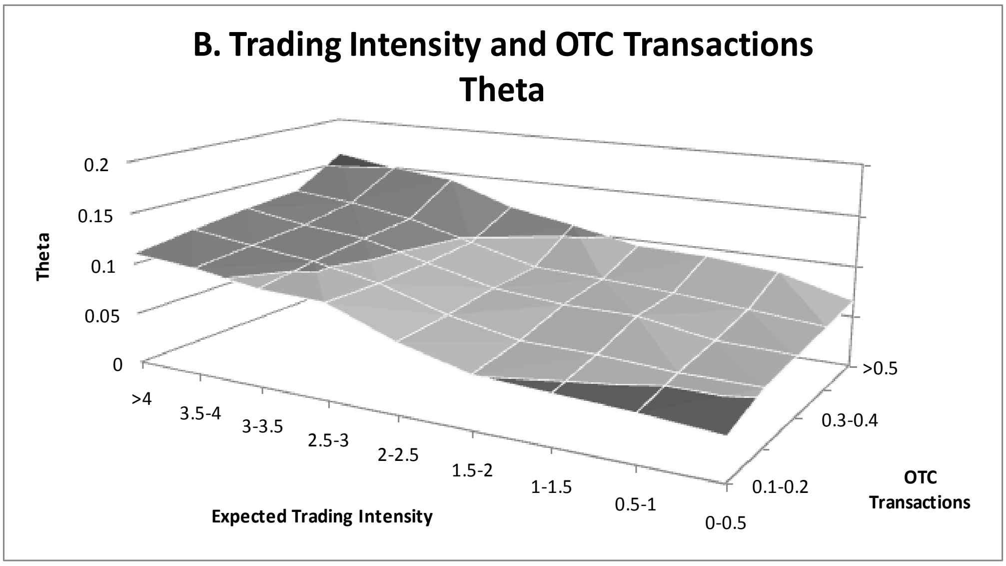


Figure 6.30: NP II Theta across Trading Intensity, Risk and OTC Transactions





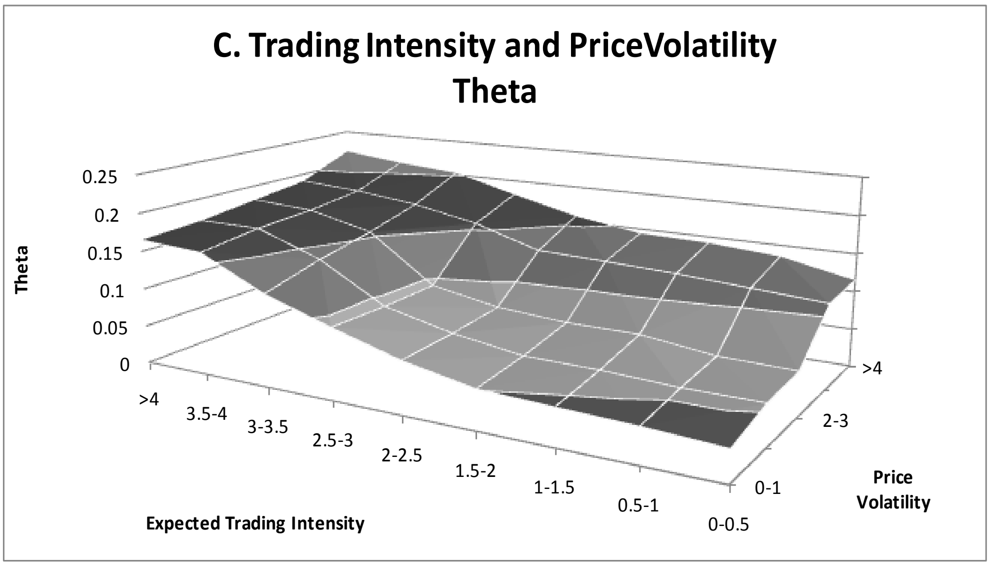
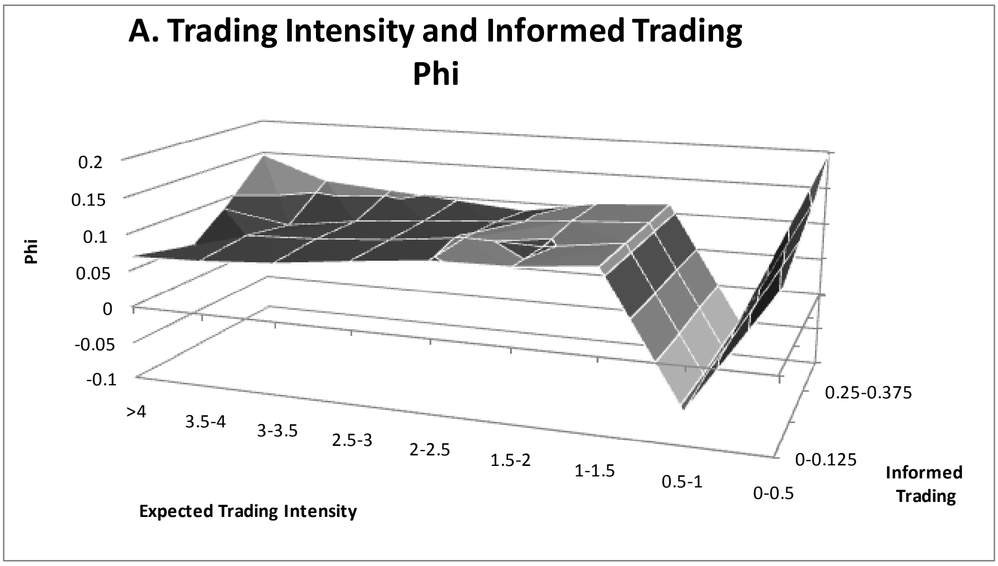
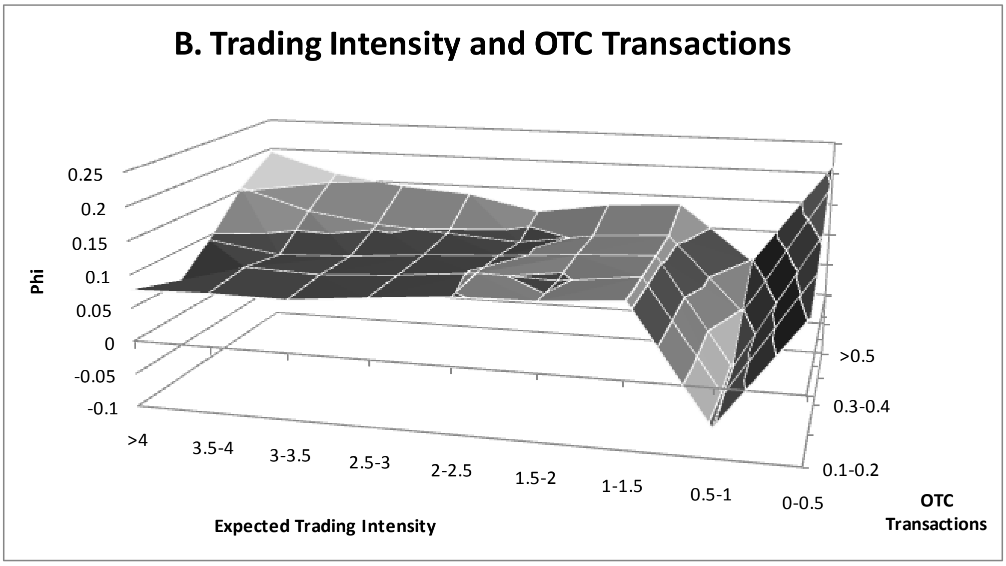


Figure 6.31: NP II Phi across Trading Intensity, Risk and OTC Transactions





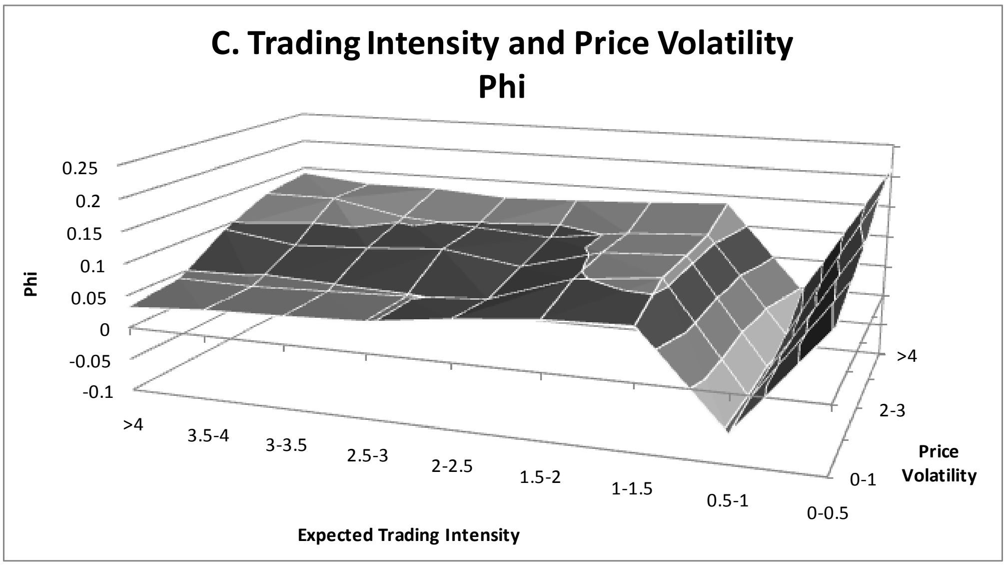


Figure 6.32: NP II Marginal Variance over Spread change across Trading Intensity, Risk and OTC Transactions